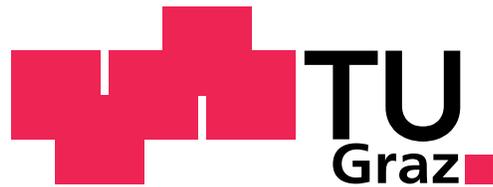


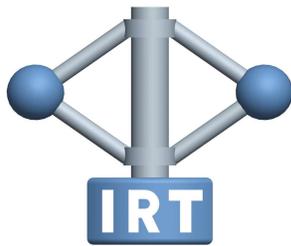
Steuerbarkeitsmaße für lineare zeitinvariante Systeme

Dipl.-Ing. (FH) Sonja Moschik



Dissertation
zur Erlangung des akademischen Grades
Doktorin der technischen Wissenschaften
an der Fakultät für Elektro- und Informationstechnik
der technischen Universität Graz

Graz, im Mai 2011



Institut für Regelungs- und Automatisierungstechnik

Technische Universität Graz
8010 Graz, Kopernikugasse 24/II

Vorstand: O.Univ.Prof Dr.techn. N. Dourdoumas

Eidesstattliche Erklärung

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbstständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt und die den benutzten Quellen wörtlich und inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

Graz, am

Vorwort

Vorliegende Arbeit entstand während meiner Tätigkeit als wissenschaftliche Assistentin am Institut für Regelungs- und Automatisierungstechnik der TU Graz.

Mein besonderer Dank gilt den Herren Professoren Nicolaos Dourdoumas, Martin Horn und Christian Magele. Ihre fachliche Unterstützung, die zahlreichen Gespräche und Diskussionen sowie manch kritische Bemerkung haben maßgeblich zum Gelingen dieser Arbeit beigetragen.

Kurzfassung

Steuerbarkeit bzw. Beobachtbarkeit stellen fundamentale Konzepte bei der Untersuchung von linearen, zeitinvarianten Systemen dar. Es existiert eine Reihe von Kriterien zur Überprüfung dieser Systemeigenschaften. Die Anwendung dieser Kriterien liefert allerdings nur eine Ja-Nein Aussage. Dadurch wird lediglich eine qualitative Charakterisierung des betrachteten Systems wiedergegeben.

In den vergangenen 50 Jahren wurde eine Vielzahl unterschiedlicher Maßzahlen zur Quantifizierung der Begriffe Steuer- und Beobachtbarkeit eingeführt. Bei der Formulierung eines Maßes, welches die Steuerbarkeit geeignet quantifiziert, stellt sich die zentrale Frage, wie die Begriffe „gut“ bzw. „schlecht“ steuerbar zu verstehen und wie diese qualitativen Begriffe in die Sprache der Mathematik zu übersetzen sind. Hierfür gibt es mehrere Möglichkeiten. Die daraus entstehenden Definitionen der Maßzahlen werden zusammengefasst, deren Eigenschaften einer mathematischen Untersuchung und einer kritischen Bewertung unterzogen. Die Analyse der Vor- und Nachteile der einzelnen Definitionen und deren Vergleich bringen nützliche und lehrreiche Einblicke in das „Wesen der Steuerbarkeit“. Ein anderer, etwas pragmatischerer Blickwinkel für eine solche Analyse ist jener aus der Sicht der Anwendung. Es werden unterschiedlichste Anwendungsgebiete der Maße vorgestellt und anhand von ausgewählten Beispielen wird überprüft, welche Maße sich in der Anwendung als effizient und zuverlässig erweisen.

Abstract

Controllability and observability are important properties of linear time-invariant systems. These properties play a crucial role in many control problems. Usually criteria for controllability and observability only provide binary results. This means that such an algebraic criterion for controllability tells us whether a system is controllable or not. It is of particular interest to know how close a controllable system is to an uncontrollable one. To this end, several measures to quantify these properties were introduced in the past 50 years. To define a measure of controllability it is important to clarify in advance what “well-controllable” or “badly-controllable” means. For this various possibilities exist. The resulting definitions of the measures are summarised and are subjected to a mathematical examination and a critical evaluation. This analysis provides an interesting insight into the quality of controll- and observability. Moreover several areas of application are introduced and the efficiency and reliability of these measures are verified based on selected examples.

Inhaltsverzeichnis

1	Einleitung	1
2	Modal- und Energiemaße	5
2.1	Modalform	5
2.1.1	Definitionen von Steuerbarkeitsmaßen	7
2.2	Übertragungsmatrix des modaltransformierten Systems	7
2.2.1	Maßdefinitionen	8
2.3	Minimale Energie	9
2.3.1	Steuerbarkeitsmaße nach KALMAN	10
2.3.2	Steuerbarkeitsmaße nach BENNINGER und RIVOIR	10
2.4	Vergleich der Maße	15
2.4.1	Mögliche Erweiterung	19
2.4.2	Diskussion der Ergebnisse	20
3	Distanzmaß	23
3.1	Analytische Lösung	26
3.1.1	Systeme 1. Ordnung	26
3.1.2	Systeme 2. Ordnung	27
3.1.3	Systeme höherer Ordnung	32
3.2	Numerische Bestimmung	34
3.2.1	Numerische Bestimmung der Größe μ_C	36
3.2.2	Numerische Bestimmung der Größe μ_R	51
3.2.3	Einfache Methoden zur Abschätzung von μ	55
3.2.4	Überblick über die Methoden zur Ermittlung von μ	56
3.2.5	Überprüfung der Forderungen	56
3.2.6	Berechnungsdauer der Algorithmen	59
3.2.7	Diskussion der Ergebnisse	62
4	Anwendungen und Beispiele	65
4.1	Beispiele	65
4.1.1	Steuerbarkeitsnormalform	65
4.1.2	Mehrgrößensystem	69
4.2	Elektromechanische Beispiele	71
4.2.1	Wagen mit Stab	72
4.2.2	Hubmagnet	79
4.2.3	Doppelpendel	82

4.2.4	Zusammenfassung der Beispiele	88
4.3	Anwendungen	91
4.3.1	Zustandsreglerentwurf	92
4.3.2	Sensorplatzierung zur Fehlerdiagnose	95
4.3.3	Bestimmung des optimalen Eingangsvektors	99
5	Anwendung der Maße auf zeitdiskrete Systeme	105
5.1	Diskretisierungsmethode und Maßzahlen	105
5.2	Maße in Abhängigkeit der Abtastzeit	106
5.2.1	Systeme 1. Ordnung	107
5.2.2	Systeme 2. Ordnung	109
5.2.3	Systeme höherer Ordnung	120
5.3	Anwendung und Deutung der Maße	121
5.3.1	Zustandsregler	122
5.3.2	Anwendbarkeit	126
5.4	Zusammenfassung	128
6	Zusammenfassung und Ausblick	131
6.1	Eingangs- Ausgangssteuerbarkeit	133
6.2	Perspektivische Beobachtbarkeit	133
	Anhang	135
A	Definitionen und Erläuterungen	135
A.1	Pseudospektrum	135
A.2	KALMAN-Zerlegung	136
A.3	KRONECKER kanonische Form	137
A.4	Störungstheorie	140
A.4.1	Das BAUER-FIKE-Theorem	140
A.4.2	Das WIELANDT-HOFFMANN-Theorem und seine Konsequenzen	141
B	Parameterwerte	143
	Literaturverzeichnis	148

Kapitel 1

Einleitung

Die Begriffe Steuerbarkeit und Beobachtbarkeit sind von fundamentaler Bedeutung bei der Analyse und Synthese zeitkontinuierlicher, linearer, zeitinvarianter Systeme (LZI-Systeme)

$$(\mathbf{A}, \mathbf{B}, \mathbf{C}) : \begin{aligned} \frac{dx}{dt} &= \mathbf{A}x + \mathbf{B}u & \mathbf{x}(0) &= \mathbf{x}_0 \\ \mathbf{y} &= \mathbf{C}x \end{aligned} \quad (1.1)$$

mit $\mathbf{A} \in \mathbb{R}^{n \times n}$, $\mathbf{B} \in \mathbb{R}^{n \times m}$ und $\mathbf{C} \in \mathbb{R}^{p \times n}$. Für die Probleme Polvorgabe, Minimalrealisierung, Entwurf optimaler Systeme und flachheitsbasierte Steuerung ist die Steuer- bzw. Beobachtbarkeit Voraussetzung für die Existenz der Lösung. Ist ein System steuerbar, dann existiert eine Eingangsfunktion u derart, dass ausgehend von einem beliebigen Anfangszustand \mathbf{x}_0 ein beliebiger Endzustand in endlicher Zeit erreicht werden kann. Ist ein System beobachtbar, so kann der unbekannte Anfangszustand \mathbf{x}_0 auf Grund des Verlaufs der Aus- und Eingangsgrößen y und u in einem endlichen Zeitintervall ermittelt werden.

Auf Grund der Dualität zweier LZI-Systeme existiert zwischen der Steuer- und der Beobachtbarkeit folgender Zusammenhang: Das System $(\mathbf{A}, \mathbf{B}, \mathbf{C})$ ist genau dann steuerbar (beobachtbar), wenn das duale System $(-\mathbf{A}^T, \mathbf{C}^T, \mathbf{B}^T)$ beobachtbar (steuerbar) ist. Deshalb wird in dieser Arbeit meist nur die Steuerbarkeit behandelt.

Für das Verständnis des Steuerbarkeitsbegriffs ist folgende Tatsache von Bedeutung:

Theorem 1 (LEE und MARKUS [28]): Betrachtet man ein steuerbares System (\mathbf{A}, \mathbf{B}) , so existiert eine positive Konstante ϵ_s derart, dass *jedes* System $(\hat{\mathbf{A}}, \hat{\mathbf{B}})$ mit der Eigenschaft

$$\|\mathbf{A} - \hat{\mathbf{A}}\| < \epsilon_s \quad \text{und} \quad \|\mathbf{B} - \hat{\mathbf{B}}\| < \epsilon_s$$

ebenfalls steuerbar ist. Umgekehrt, falls ein System $(\tilde{\mathbf{A}}, \tilde{\mathbf{B}})$ nicht steuerbar ist, so existiert zu *jeder* positiven Konstante ϵ ein steuerbares System (\mathbf{A}, \mathbf{B}) mit der Eigenschaft

$$\|\mathbf{A} - \tilde{\mathbf{A}}\| < \epsilon \quad \text{und} \quad \|\mathbf{B} - \tilde{\mathbf{B}}\| < \epsilon.$$

Dieses Theorem sagt somit aus, dass die Menge der steuerbaren LZI-Systeme offen und dicht ist bzw. dass es in der Menge der LZI-Systeme nur an singulären Punkten nicht steuerbare Systeme gibt. Mit anderen Worten: Die Steuerbarkeit ist eine generische Eigenschaft, man kann sagen, dass ein LZI-System im „Normalfall“ steuerbar ist.

Es existiert eine Reihe von einfach anzuwendenden Kriterien zur Untersuchung der Steuer- und

Beobachtbarkeit von LZI-Systemen. Diese Kriterien liefern im Allgemeinen nur eine Ja-Nein Aussage. Das heißt, man erhält die Information, ob ein System steuerbar bzw. beobachtbar ist oder nicht. Neben dem Vorhanden- oder nicht Vorhandensein dieser Systemeigenschaften stellen sich aber weitere interessante Fragen: Wie „leicht“ ist die Strecke überhaupt zu steuern? Wo sollen Aktoren und Sensoren angebracht werden, damit die Strecke „gut“ steuerbar ist? Welche Zustandsvariablen sollen gemessen werden um ein „gute“ Beobachtbarkeit zu erzielen? Wie groß ist die benötigte Stellgröße bzw. wieviel Energie ist notwendig, um den gewünschten Zustand zu erreichen? Wie kann das System verändert werden, um die benötigte Steuerenergie zu minimieren?

Die vorhandenen, einfach strukturierten Kriterien liefern bei der numerischen Auswertung in manchen Fällen falsche Aussagen. Im Folgenden wird dies für das Kriterium nach KALMAN bzw. für jenes nach HAUTUS gezeigt.

Das KALMAN-Kriterium besagt, dass ein System (\mathbf{A}, \mathbf{B}) genau dann steuerbar ist, wenn die rechteckige (n, nm) -Steuerbarkeitsmatrix

$$\mathbf{S}_u = [\mathbf{B}, \mathbf{A}\mathbf{B}, \mathbf{A}^2\mathbf{B}, \dots, \mathbf{A}^{n-1}\mathbf{B}]$$

den Rang n besitzt.

Nach HAUTUS ist das System (\mathbf{A}, \mathbf{B}) genau dann steuerbar, wenn die rechteckige $(n, n + m)$ -Matrix

$$\mathbf{H}_u = [\lambda\mathbf{E} - \mathbf{A}, \mathbf{B}]$$

für alle Eigenwerte $\lambda = \lambda_i$ ($i = 1, \dots, n$) der Matrix \mathbf{A} den Rang n hat.

Sowohl für das KALMAN- als auch für das HAUTUS-Kriterium ist die Rangbestimmung einer Matrix notwendig. Der Rang kann durch eine Singulärwertzerlegung der Matrix bestimmt werden. Hierfür gilt folgender Satz:

Lemma 2: Es seien

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n$$

die geordneten Singulärwerte einer (n, l) -Matrix \mathbf{M} . Genau dann, wenn

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0 \text{ und } \sigma_{r+1} = \dots = \sigma_n = 0$$

gilt, hat die Matrix \mathbf{M} den Rang r .

Die Rangbestimmung mittels Singulärwertzerlegung ist zuverlässig, da die Singulärwerte nicht sehr empfindlich auf Änderungen der Matrixeinträge reagieren [38]. Trotzdem kann die Rangbestimmung im Rahmen der numerischen Ermittlung fehlerhaft ausfallen.

Wir betrachten das aus [48] entnommene Beispiel mit der Systemmatrix in Dreieckstruktur

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & \dots & \dots & 0 \\ 20 & 2 & & & \vdots \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & & 20 & 19 & 0 \\ 0 & \dots & 0 & 20 & 20 \end{bmatrix}$$

und den beiden möglichen Eingangsvektoren

$$\mathbf{B}_1 = \begin{bmatrix} 20 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \text{bzw.} \quad \mathbf{B}_2 = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 20 \end{bmatrix}.$$

Es kann analytisch bestimmt werden, dass das System $(\mathbf{A}, \mathbf{B}_1)$ steuerbar ist, da \mathbf{B}_1 Komponenten in alle Eigenrichtungen von \mathbf{A} besitzt. Das System $(\mathbf{A}, \mathbf{B}_2)$ ist nicht steuerbar, da \mathbf{B}_2 ein Eigenvektor von \mathbf{A} ist.

Durch Singulärwertzerlegung der Steuerbarkeitsmatrix \mathbf{S}_u des Systems $(\mathbf{A}, \mathbf{B}_1)$ erhält man numerisch Singulärwerte gleich Null. Daraus wird die falsche Schlussfolgerung gezogen, dass \mathbf{S}_u nicht den vollen Rang hat und somit das System $(\mathbf{A}, \mathbf{B}_1)$ nicht steuerbar ist. Der Grund für diese numerischen Probleme liegt darin, dass die Matrix \mathbf{S}_u sowohl sehr kleine als auch Elemente der Größenordnung 20^{19} enthält.

Obige Matrix \mathbf{A} besitzt die Eigenwerte $\lambda_i = i$ ($i = 1, \dots, 20$), die schlecht konditioniert sind. D.h., modifiziert man die Matrix \mathbf{A} , indem man das Element $a_{1,20} = 10^{-12}$ statt Null wählt, erhält man eine Matrix $\tilde{\mathbf{A}}$ mit den Eigenwerten

$$\begin{array}{ll} \tilde{\lambda}_1 = & 1 & \tilde{\lambda}_{12,13} = & 12,6260 \pm j0,9150 \\ \tilde{\lambda}_2 = & 2,0008 & \tilde{\lambda}_{14,15} = & 14,6438 \pm j0,4678 \\ \tilde{\lambda}_3 = & 2,9927 & \tilde{\lambda}_{16} = & 16,1443 \\ \tilde{\lambda}_4 = & 4,0449 & \tilde{\lambda}_{17} = & 16,9551 \\ \tilde{\lambda}_5 = & 4,8557 & \tilde{\lambda}_{18} = & 18,0073 \\ \tilde{\lambda}_{6,7} = & 6,3562 \pm j0,4678 & \tilde{\lambda}_{19} = & 18,9992 \\ \tilde{\lambda}_{8,9} = & 8,3740 \pm j0,9150 & \tilde{\lambda}_{20} = & 20,0000. \\ \tilde{\lambda}_{10,11} = & 10,5000 \pm j1,0686 & & \end{array}$$

Fast alle der obigen Eigenwerte unterscheiden sich nicht nur betragsmäßig von den ursprünglichen, sondern man erhält durch die „Störung“ konjugiert komplexe Eigenwerte, während die Originalmatrix nur reelle Eigenwerte besitzt.

Das Vorhandensein schlecht konditionierter Eigenwerte erklärt, warum die numerische Überprüfung der Steuerbarkeit mit Hilfe des HAUTUS-Kriteriums in manchen Fällen zu falschen Ergebnissen führen kann. Numerisch stabile Algorithmen zur Bestimmung der Eigenwerte λ_i einer gegebenen Matrix \mathbf{A} liefern die Eigenwerte einer in der Nähe liegenden Matrix $\tilde{\mathbf{A}}$ [38], die sich gerade im Falle schlecht konditionierter Eigenwerte stark von den ursprünglichen unterscheiden können. Angenommen, der verwendete Algorithmus liefert wegen numerischer Fehler die „falschen“ Eigenwerte $\tilde{\lambda}_i$. Wird nun zur Überprüfung der Steuerbarkeit der Rang von $[\tilde{\lambda}_i \mathbf{E} - \mathbf{A}, \mathbf{B}_2]$ berechnet, so ist dieser gleich n für alle Eigenwerte $\tilde{\lambda}_i$. Daraus wird die falsche Schlussfolgerung gezogen, dass das System $(\mathbf{A}, \mathbf{B}_2)$ steuerbar ist.

Auf die bei der Auswertung von Steuerbarkeitskriterien auftretenden numerischen Schwierigkeiten (Rang- sowie Eigenwertbestimmung) wird z. B. in [38] näher eingegangen.

Die Tatsachen, dass ein LZI-System im Normalfall steuerbar ist, die bekannten Kriterien zur Überprüfung der Steuerbarkeit nur eine Ja-Nein Aussage und in manchen Fällen falsche Ergebnisse liefern sind nur einige der Gründe, die die Einführung von Maßzahlen zur Quantifizierung der Steuerbarkeit motivieren. Zur Bewertung der Steuer- und Beobachtbarkeit linearer

Systeme wurden in der Vergangenheit eine Vielzahl unterschiedlicher Maßzahlen eingeführt. Diese können auf Grund der Art der Quantifizierung und damit wegen ihrer prinzipiellen Aussage über die Steuerbarkeit in drei Klassen eingeteilt werden: *Modalmaß*, *Energiemaß* und *Distanzmaß*.

Im ersten Teil dieser Arbeit werden in den Abschnitten 2.1 - 2.3 die wichtigsten Definition der Modal- und Energiemaße vorgestellt. Auf Grund der Ähnlichkeit der beiden Klassen werden diese gemeinsam in Kapitel 2 behandelt und in Abschnitt 2.4 direkt miteinander verglichen. Die Maße dieser beiden Klassen ermöglichen, eine Information darüber zu erhalten, wie „gut“ oder „schlecht“ eine Zustandsvariable steuerbar bzw. beobachtbar ist. D.h. diese Maßzahlen bewerten die Steuer- und/oder Beobachtbarkeit *einzelner* Zustandsvariablen. Damit können dann Aussagen darüber getroffen werden, welche Zustandsvariablen gemessen werden müssen, um eine hinreichend gute Beobachtbarkeit zu erreichen. Hat man bei der Wahl der Platzierung der Aktoren und Sensoren gewisse Freiheiten, kann mit Hilfe der Maße versucht werden, diese möglichst effektiv anzuordnen. Des Weiteren kann man mit Hilfe der Maße eine Aussage über die Empfindlichkeit des Systems bezüglich Parameterungenauigkeiten treffen. D.h., wenn das nicht exakt modellierte System beispielsweise „schlecht“ steuerbar ist, ist unter Umständen das reale System auf Grund von Parametertoleranzen oder Parameterschwankungen nicht steuerbar. Mit Hilfe der Maßzahlen kann die Dominanz einzelner Systemteile beurteilt werden. Sie ermöglicht, die Ordnung eines komplexen Systems durch Vernachlässigung der Systemteile mit geringem Einfluss auf das Gesamtsystem zu reduzieren.

Auf Grund der erwähnten numerischen Schwierigkeiten wurde von PAIGE in [38] für eine zuverlässige Beurteilung der Steuerbarkeit die Maßzahl μ eingeführt. Diese Maßzahl liefert eine Aussage über die „Größe“ der Störung, die notwendig ist, damit ein steuerbares System nicht steuerbar wird. Ist μ größer als ein bestimmter Wert μ_0 , dann ist das System mit Sicherheit steuerbar. Der Wert μ_0 wird aus den Unsicherheiten der Systemmatrizen und einer vorgegebenen Toleranzschranke für die numerischen Fehler berechnet. Das Theorem nach LEE und MARKUS lässt folgende Interpretation zu: Der Wert μ kann als der kleinste Abstand zu einem nicht steuerbaren System interpretiert werden und wird aus diesem Grund als Distanzmaß bezeichnet. Kapitel 3 beschäftigt sich mit dem Distanzmaß μ und dessen Berechnungsmöglichkeiten. Dieses Maß liefert eine Aussage über die Güte der Steuerbarkeit des *gesamten* Systems.

Anhand einfacher Beispiele werden die Schwierigkeiten, die bei der Anwendung der verschiedenen Maße auftreten können, aufgezeigt.

Im zweiten Teil werden die vorgestellten Maße miteinander verglichen und anhand von Beispielen auf deren Anwendbarkeit überprüft. Mögliche Anwendungsgebiete für die Maßzahlen werden in Abschnitt 4.3 genauer erläutert. Kapitel 5 beinhaltet die Anwendung der Maße auf zeitdiskrete, lineare, zeitinvariante Systeme.

An dieser Stelle sei erwähnt, dass die Steuerbarkeit eine notwendige und hinreichende Bedingung für die Flachheit von LZI-Systemen ist. Diese beiden Systemeigenschaften sind für LZI-Systeme somit äquivalent. Für diesen Fall können die Steuerbarkeitsmaße damit auch als Maß für die Flachheit eines Systems verstanden und die in dieser Arbeit gewonnenen Erkenntnisse auch auf diese Systemeigenschaft übertragen werden.

Kapitel 2

Modal- und Energiemaße

Will man die Steuerbarkeit eines Systems quantifizieren, muss zunächst geklärt werden, wie die Begriffe „gut“ und „schlecht“ steuerbar zu verstehen sind. Um einen besseren Einblick zu erhalten, betrachten wir zunächst ein zeitkontinuierliches LZI-System mit einer besonderen Struktur, der sogenannten Modalform. Die auf der Modalform basierenden Maße werden der Klasse der Modalmaße zugeordnet.

2.1 Modalform

Die Modalform kann aus Gleichung (1.1) durch die Zustandstransformation

$$\mathbf{x} = [\mathbf{x}_{R1}, \mathbf{x}_{R2}, \dots, \mathbf{x}_{Rn}] \mathbf{z} = \mathbf{X}_R \mathbf{z}$$

erhalten werden. Die Spalten der Matrix \mathbf{X}_R entsprechen Rechts-Eigenvektoren \mathbf{x}_{Ri} der Matrix \mathbf{A} . Unter der Voraussetzung, dass die (n, n) -Matrix \mathbf{A} verschiedene Eigenwerte λ_i ($i = 1, \dots, n$) besitzt, sind die zugehörigen Eigenvektoren linear unabhängig und die Transformationsmatrix \mathbf{X}_R ist regulär. Mit der Beziehung

$$\mathbf{X}_R^{-1} = \mathbf{X}_L,$$

wobei die Zeilen von \mathbf{X}_L Links-Eigenvektoren \mathbf{x}_{Li} der Matrix \mathbf{A} entsprechen, erhält man das in Modalform transformierte System

$$\begin{aligned} \frac{d\mathbf{z}}{dt} &= \mathbf{X}_L \mathbf{A} \mathbf{X}_R \mathbf{z} + \mathbf{X}_L \mathbf{B} \mathbf{u} \\ \mathbf{y} &= \mathbf{C} \mathbf{X}_R \mathbf{z} \end{aligned} \tag{2.1}$$

mit der nun in Diagonalform vorliegenden Systemmatrix

$$\hat{\mathbf{A}} := \mathbf{X}_L \mathbf{A} \mathbf{X}_R = \text{diag}(\lambda_i) = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{bmatrix}.$$

Die Ein- und Ausgangsmatrizen des modaltransformierten Systems erhält man durch die Beziehungen

$$\hat{\mathbf{B}} := \mathbf{X}_L \mathbf{B} = \begin{bmatrix} \hat{\mathbf{b}}_1^T \\ \hat{\mathbf{b}}_2^T \\ \vdots \\ \hat{\mathbf{b}}_n^T \end{bmatrix} \quad \text{und} \quad \hat{\mathbf{C}} := \mathbf{C} \mathbf{X}_R = [\hat{\mathbf{c}}_1 \quad \hat{\mathbf{c}}_2 \quad \dots \quad \hat{\mathbf{c}}_n].$$

Relevante Eigenschaften des Systems, wie z.B. Stabilität, Steuer- und Beobachtbarkeit, bleiben bei einer regulären Zustandstransformation erhalten.

Die Transformation auf Modalform führt zu einem *entkoppelten* System mit n Differentialgleichungen der Form

$$\frac{dz_i}{dt} = \lambda_i z_i + \hat{\mathbf{b}}_i^T \mathbf{u}.$$

Der Einfluss der Zeilen $\hat{\mathbf{b}}_i^T$ der transformierten Eingangsmatrix $\hat{\mathbf{B}}$ auf die Steuerbarkeit bzw. auf die Beeinflussbarkeit der entsprechenden Teilsysteme ist nun direkt sichtbar, wie Abbildung 2.1 zeigt.

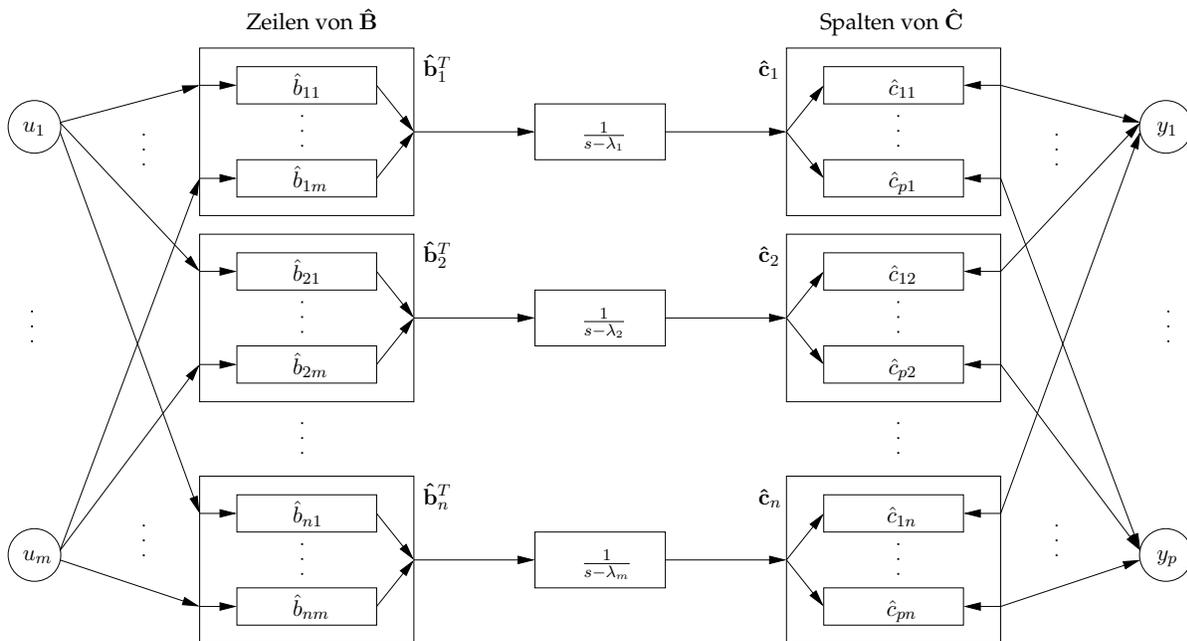


Abbildung 2.1: Modalform [39]

Dies führt direkt zum

Theorem 3 (Kriterium nach GILBERT): Ein System der Ordnung n in Modalform ist bei einfachen Eigenwerten λ_i ($i = 1, \dots, n$) genau dann steuerbar, wenn die Eingangsmatrix $\hat{\mathbf{B}}$ des transformierten Systems keine Nullzeile besitzt.

2.1.1 Definitionen von Steuerbarkeitsmaßen

Die Steuerbarkeitsmaße beruhen auf der Idee, dass kleine Einträge in den entsprechenden Zeilen $\hat{\mathbf{b}}_k^T$ der Eingangsmatrix $\hat{\mathbf{B}}$ auf eine „schlecht“ steuerbare Zustandsvariable hinweisen. Es werden Aussagen im sogenannten Modalraum getroffen, weshalb in der Literatur oft von der Steuerbarkeit eines Eigenwertes gesprochen wird, da sich die Zustandsvariablen z_i direkt einem Eigenwert λ_i zuordnen lassen. In der Vergangenheit wurden unter Benutzung der Zeilen von $\hat{\mathbf{B}}$ Maßzahlen definiert, die im Folgenden kurz zusammengefasst werden.

KONNO [25]: Das modale Steuerbarkeitsmaß ergibt sich durch Normieren der Zeilen der transformierten Eingangsmatrix $\hat{\mathbf{B}}$ zu

$$a_i = \sqrt{\hat{\mathbf{b}}_i^H \hat{\mathbf{b}}_i}. \quad (2.2)$$

LÜCKEL und MÜLLER [27]: Das Maß

$$\kappa_{si} = \frac{\mathbf{x}_{Li}^H \mathbf{B} \mathbf{B}^T \mathbf{x}_{Li}}{\mathbf{x}_{Li}^H \mathbf{x}_{Li}} e^{-2\text{Re}\{\lambda_i\}} = \frac{\hat{\mathbf{b}}_i^H \hat{\mathbf{b}}_i}{\mathbf{x}_{Li}^H \mathbf{x}_{Li}} e^{-2\text{Re}\{\lambda_k\}} \quad (2.3)$$

beinhaltet einerseits die Zeilen der Eingangsmatrix $\hat{\mathbf{B}}$ und andererseits den Term $e^{-2\text{Re}\{\lambda_i\}}$. Dieser bewirkt eine unterschiedliche Bewertung von stabilen und instabilen Eigenwerten. Die zu Grunde liegende Idee für die Einführung dieses Terms ist, dass ein stabiler Eigenwert den Regelvorgang in den Ursprung unterstützt, während ein instabiler Eigenwert diesem entgegenwirkt.

LITZ [32]: Da die von LÜCKEL und MÜLLER eingeführte Verkoppelung zwischen Steuerbarkeit und Stabilität zu unerwünschten Verzerrungen der Maßzahlen führen kann, wie später gezeigt wird, wurde das von der Stabilitätsverkoppelung bereinigte Maß

$$\kappa_i = \frac{\mathbf{x}_{Li}^H \mathbf{B} \mathbf{B}^T \mathbf{x}_{Li}}{\mathbf{x}_{Li}^H \mathbf{x}_{Li}} \quad (2.4)$$

eingeführt.

2.2 Übertragungsmatrix des modaltransformierten Systems

Die Übertragungsmatrix des Systems (2.1) errechnet sich unmittelbar aus

$$\begin{aligned} \mathbf{G}(s) &= \hat{\mathbf{C}}(s\mathbf{E} - \hat{\mathbf{A}})^{-1} \hat{\mathbf{B}} \\ &= [\hat{\mathbf{c}}_1 \quad \hat{\mathbf{c}}_2 \quad \dots \quad \hat{\mathbf{c}}_n] \begin{bmatrix} \frac{1}{s-\lambda_1} & 0 & \dots & 0 \\ 0 & \frac{1}{s-\lambda_2} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \frac{1}{s-\lambda_n} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{b}}_1^T \\ \hat{\mathbf{b}}_2^T \\ \vdots \\ \hat{\mathbf{b}}_n^T \end{bmatrix} \\ &= \sum_{i=1}^n \hat{\mathbf{c}}_i \hat{\mathbf{b}}_i^T \frac{1}{s - \lambda_i}. \end{aligned}$$

Die Übertragungsmatrix $\mathbf{G}_i(s)$ eines entkoppelten Teilsystems mit dem Eigenwert λ_i lautet

$$\begin{bmatrix} y_1 \\ \vdots \\ y_p \end{bmatrix} = \frac{1}{s - \lambda_i} \underbrace{\begin{bmatrix} \hat{c}_{1i}\hat{b}_{i1} & \cdots & \hat{c}_{1i}\hat{b}_{im} \\ \vdots & \ddots & \vdots \\ \hat{c}_{pi}\hat{b}_{i1} & \cdots & \hat{c}_{pi}\hat{b}_{im} \end{bmatrix}}_{\mathbf{G}_i(s)} \begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix}.$$

Es wird ersichtlich, dass das Produkt aus \hat{c}_{ki} und \hat{b}_{ij} für die Steuer- und Beobachtbarkeit von Bedeutung ist. Wird nämlich die Matrix \mathbf{G}_i zur Nullmatrix, dann ist dieses Teilsystem nicht steuerbar oder nicht beobachtbar bzw. weder steuerbar noch beobachtbar. Wenn eine Spalte $\hat{c}_i\hat{b}_{ij}$ dieser Matrix dem Nullvektor entspricht, hat die zugehörige Eingangsgröße u_j keinen Einfluss auf die Ausgangsgrößen. Enthält die Matrix \mathbf{G}_i eine Nullzeile $\hat{\mathbf{b}}_i^T\hat{c}_{ji}$, dann enthält die zugehörige Ausgangsgröße y_j keinerlei Information.

2.2.1 Maßdefinitionen

Auch hier kann wieder die Überlegung angestellt werden, dass „kleine“ Werte von $\hat{c}_{ki}\hat{b}_{ij}$ auf „schlechte“ Steuerbarkeit und/oder Beobachtbarkeit hinweisen. Die im Folgenden vorgestellten Maße für Regelbarkeit, Polverschiebbarkeit und Dominanz wurden auf Grund unterschiedlicher Überlegungen eingeführt, wodurch verschiedene Interpretationen zustandekamen. Diese Maße bemessen, unter Benutzung des Produkts $\hat{c}_{ki}\hat{b}_{ij}$, die gleiche Systemeigenschaft und liefern somit, wie auch anhand der Maßdefinitionen leicht erkennbar, ähnliche Ergebnisse.

Regelbarkeitsmaß nach HIPPE [21]: Das Regelbarkeitsmaß gibt Auskunft über die Beeinflussbarkeit eines Eigenwerts durch einen Regler. Beim Entwurf eines optimalen Zustandsreglers zeigt sich, dass u.U. einige Eigenwerte der Strecke relativ stark verändert werden, während sich andere nur sehr wenig von der ursprünglichen Lage entfernen. Letztere werden in [21] als schlecht regelbar bezeichnet. Das die Regelbarkeit bewertende Maß für den Eigenwert λ_i , bezogen auf das Eingangs-Ausgangspaar (u_j, y_k) , ist definiert zu

$$|\rho_i^{kj}| = |\hat{c}_{ki}\hat{b}_{ij}|. \quad (2.5)$$

Polempfindlichkeitsmaß nach LITZ [30]: Bei konstanter Rückführung des Systemausgangs auf den Systemeingang

$$\mathbf{u} = \mathbf{R}\mathbf{y} = \begin{bmatrix} r_{11} & \cdots & r_{1p} \\ \vdots & \ddots & \vdots \\ r_{m1} & \cdots & r_{mp} \end{bmatrix} \mathbf{y}$$

ist das *absolute* Polempfindlichkeitsmaß durch

$$S_{r_{jk}}^{\lambda_i} = \frac{\partial \lambda_i}{\partial r_{jk}} = \hat{c}_{ki}\hat{b}_{ij} \quad (2.6)$$

und das *relative* Polempfindlichkeitsmaß durch

$$s_{r_{jk}}^{\lambda_i} = \frac{\partial \lambda_i}{\partial r_{jk}} \frac{1}{\lambda_i} = \frac{\hat{c}_{ki}\hat{b}_{ij}}{\lambda_i} \quad (2.7)$$

mit λ_i als Eigenwert des rückgekoppelten Systems definiert. Beide Polempfindlichkeitsmaße können im Allgemeinen komplexe Werte annehmen. Hierbei entspricht $\operatorname{Re}\{S_{r_{jk}}^{\lambda_i}\}$ bzw. $\operatorname{Re}\{s_{r_{jk}}^{\lambda_i}\}$ der Empfindlichkeit des Realteils und $\operatorname{Im}\{S_{r_{jk}}^{\lambda_i}\}$ bzw. $\operatorname{Im}\{s_{r_{jk}}^{\lambda_i}\}$ der Empfindlichkeit des Imaginärteils des Eigenwerts λ_i .

Dominanzmaß nach LITZ [32, 29]: Das Maximalmaß M_i und das Summenmaß Σ_i , bezogen auf den Eigenwert λ_i , beruhen auf der Untersuchung der Antwort der k -ten Zustandsvariable bei Erregung des j -ten Eingangs durch eine Sprungfunktion. Die Antwort hängt in einfacher Weise von Faktoren der Form $\frac{\hat{c}_{ki}\hat{b}_{ij}}{\lambda_i}$ ab. Aus den für jeden Eigenwert λ_k vorliegenden $m \cdot p$ Faktoren werden zwei Gesamtmaße wie folgt definiert

$$M_i = \max_{k=1}^p \left(\max_{j=1}^m \left| \frac{\hat{c}_{ki}\hat{b}_{ij}}{\lambda_i} \right| \right) \quad (2.8)$$

$$\Sigma_i = \sum_{k=1}^p \sum_{j=1}^m \left| \frac{\hat{c}_{ki}\hat{b}_{ij}}{\lambda_i} \right|. \quad (2.9)$$

Das Maximalmaß trifft eine generelle Aussage über die Dominanz des Eigenwertes. Durch die Bestimmung des Summenmaßes wird der Fall erfasst, dass ein Eigenwert mit einem hohen Beitrag in einem einzigen Übertragungskanal trotz eines hohen Maximalmaßes nur einen geringen Beitrag zum Systemverhalten leistet [39].

2.3 Minimale Energie

Eine weitere Möglichkeit zur Quantifizierung der Steuerbarkeit ergibt sich, wenn ein Zustand \mathbf{x}_1 besser steuerbar als ein Zustand \mathbf{x}_2 genannt wird, falls für die Überführung nach $\mathbf{0}$ weniger Steuerenergie notwendig ist. Hierbei wird die Steuerenergie im vorgegebenen Intervall $[0, t_1]$ durch

$$W(t_1, \mathbf{u}) = \int_0^{t_1} \mathbf{u}^T(\tau) \mathbf{u}(\tau) d\tau$$

definiert. Mit der Wahl

$$\mathbf{u}(t) = -\mathbf{B}^T e^{-\mathbf{A}^T t} \mathbf{Q}_S^{-1}(t_1) \mathbf{x}_0$$

unter Verwendung der GRAMschen Steuerbarkeitsmatrix

$$\mathbf{Q}_S(t_1) := \int_0^{t_1} e^{-\mathbf{A}\tau} \mathbf{B} \mathbf{B}^T (e^{-\mathbf{A}\tau})^T d\tau \quad (2.10)$$

wird der beliebige Anfangszustand \mathbf{x}_0 eines steuerbaren Systems mit der minimal notwendigen Energie

$$W_S := \min_{\mathbf{u}} \{W(t_1, \mathbf{u})\} = \mathbf{x}_0^T \mathbf{Q}_S^{-1}(t_1) \mathbf{x}_0 \quad (2.11)$$

in den Zustand $\mathbf{x}(t_1) = \mathbf{0}$ überführt. Die minimale Energie bzw. die GRAMsche Steuerbarkeitsmatrix ist die Grundlage für die in den nachfolgenden Abschnitten angeführten Maße, weshalb diese der Klasse der Energiemaße zugeordnet werden.

2.3.1 Steuerbarkeitsmaße nach KALMAN

KALMAN schlägt als Maß für die „Güte“ der Steuerbarkeit die Spur bzw. die Determinante der Inversen GRAM-Matrix \mathbf{Q}_S^{-1} vor.

Die Determinante ist proportional zum Volumen eines Hyperellipsoids [37]

$$\text{Volumen } \{ \mathbf{x}_0^T \mathbf{Q}_S^{-1}(t_1) \mathbf{x}_0 = 1 \} \sim \sqrt{\det\{\mathbf{Q}_S(t_1)\}},$$

in dem Anfangszustände liegen, die mit einer bestimmten minimalen Energie nach Null gebracht werden können [23]. Eine mögliche Maßdefinition ist

$$d_1 = \det\{\mathbf{Q}_S^{-1}\}. \quad (2.12)$$

Die Spur gibt den Mittelwert der minimal notwendigen Steuerenergie an, um Anfangswerte, die der Beziehung

$$\|\mathbf{x}_0\| = 1$$

genügen, nach Null zu überführen. Als Maß eignet sich der mit der Systemordnung n gewichtete Kehrwert der Spur

$$d_2 = \frac{n}{\text{spur}\{\mathbf{Q}_S^{-1}\}}, \quad (2.13)$$

da dadurch für die beiden Maße d_1 und d_2 gilt, je kleiner das Maß desto „schlechter“ steuerbar ist das System [37].

Eine weitere Möglichkeit für ein Maß ist der größte Eigenwerte der Inversen GRAMschen Matrix, da dieser die maximal notwendige minimale Energie angibt:

$$\max_{\|\mathbf{x}_0\|=1} \{\mathbf{W}_S\} = \lambda_{\max}(\mathbf{Q}_S^{-1})$$

Im Gegensatz zu den in den vorherigen Abschnitten vorgestellten Maßen, bewerten diese Maße nicht die Steuerbarkeit der einzelnen Zustandsvariablen, sondern die Steuerbarkeit des Gesamtsystems und zählen somit nicht zu den Modalmaßen.

2.3.2 Steuerbarkeitsmaße nach BENNINGER und RIVOIR

Die von BENNINGER und RIVOIR in [2] definierten Maße resultieren aus der Beziehung (2.11). Die Idee besteht darin, die Energie W_S vorzugeben und die zugehörigen Auslenkungen einer Zustandsvariablen zu untersuchen. Die Menge aller Zustände, die bei einem vorgegebenen Wert W_S in den Ursprung überführt werden können, entspricht einem Hyperellipsoid, dessen Form und Lage im Zustandsraum von W_S bestimmt wird. Für $W_S = 1$ wird als Maß für die Steuerbarkeit der Zustandsvariablen x_i der Abstand zwischen Koordinatenursprung und dem Schnittpunkt des Ellipsoids mit der positiven x_i -Achse vorgeschlagen. Er ist durch

$$m_i = [(\mathbf{Q}_S^{-1})_{ii}]^{-\frac{1}{2}} \quad (2.14)$$

gegeben, wobei mit $(\mathbf{Q}_S^{-1})_{ii}$ das i -te Diagonalelement der Matrix \mathbf{Q}_S^{-1} symbolisiert wird. Ein weiteres, in [2] eingeführtes Maß bewertet die Beeinflussbarkeit einer Zustandsvariablen. Eine Zustandsvariable x_i heißt beeinflussbar, wenn es zu jedem \mathbf{x}_0 eine Eingangsfunktion $\mathbf{u}(t)$

gibt, welche die i -te Komponente des Zustandsvektors in endlicher Zeit T nach Null überführt. Zur Bestimmung des Beeinflussbarkeitsmaßes dient die maximale Auslenkung der i -ten Komponente, die durch

$$v_i = [(\mathbf{Q}_S)_{ii}]^{\frac{1}{2}} \quad (2.15)$$

berechnet wird (vgl. Abbildung 2.2).

Das Steuerbarkeitsmaß m_i gibt demnach die größtmögliche Auslenkung auf der x_i -Achse an, während das Beeinflussbarkeitsmaß v_i angibt, wie weit die i -te Komponente ausgelenkt werden kann, wenn auf die übrigen Komponenten keine Rücksicht genommen wird.

Wenn mindestens eine Zustandsvariable nicht steuerbar ist, ist die GRAMsche Matrix singulär. Um die Steuerbarkeitsmaße der übrigen Zustandsvariablen zu berechnen, wird in diesem Fall bei der Ermittlung der Inversen auf die Pseudoinverse zurückgegriffen (siehe [2]).

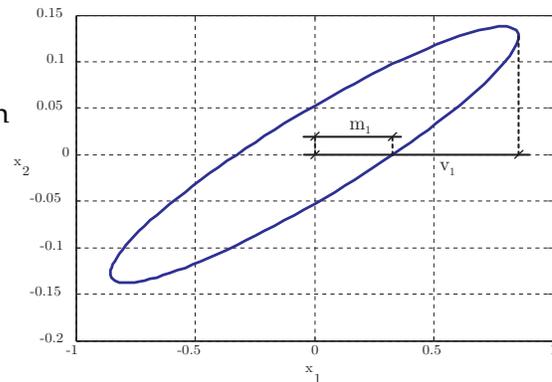


Abbildung 2.2: Bestimmung der Maßzahlen nach BENNINGER

2.3.2.1 Wahl der Endzeit t_1

In der Definitionsgleichung der GRAMschen Matrix (2.10) ist der Parameter t_1 frei wählbar, weshalb sich die Frage stellt, wie dieser möglichst sinnvoll gewählt werden kann. Folgendes Beispiel zeigt, wie sich die Wahl von t_1 auf das Steuerbarkeitsmaß auswirkt.

Beispiel 2.1: Gegeben sei das System

$$\mathbf{A} = \begin{bmatrix} \alpha & 1 \\ 0 & -2 \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

mit den Eigenwerten $\lambda_1 = \alpha$ und $\lambda_2 = -2$. In Abbildung 2.3 ist das Steuerbarkeitsmaß m_1 der Zustandsvariablen x_1 für sechs verschiedene Werte des Parameters α in Abhängigkeit der Zeit t_1 dargestellt.

Für negative Eigenwerte sieht man, dass die Wahl von t_1 einen großen Einfluss auf das Steuerbarkeitsmaß hat. Je größer t_1 wird, desto größer wird auch das Maß, da bei asymptotisch stabilen Eigenwerten mehr Zeit vorhanden ist, um den Anfangszustand ohne Einwirkung von außen in den Ursprung zu überführen. Dadurch wird die notwendige minimale Energie natürlich klein. Man sieht auch, dass je „schneller“ der Eigenwert ist (d.h. je größer $|\alpha|$ mit $\alpha < 0$), desto größer wird das Steuerbarkeitsmaß. Auf Grund dieses „schnelleren“ Eigenwerts kann der Anfangszustand auch schneller nach Null überführt werden.

Genau das Gegenteil gilt für positive Eigenwerte. In diesem Fall führen die „schnellen“ Eigenwerte zu einem kleineren Steuerbarkeitsmaß, da die Eigenwerte der Überführung in den Ursprung stärker entgegen wirken. Für einen positiven Eigenwert strebt das Maß für $t_1 \rightarrow \infty$ einem konstanten Wert zu. In diesem Fall wirkt sich die Wahl von t_1 für hinreichend große t_1 nur noch wenig auf das Steuerbarkeitsmaß aus.

Wird t_1 klein gewählt, werden die „schnellen“ Eigenwerte (betragsmäßig große Eigenwerte) bzw. wird t_1 groß gewählt, werden die „langsamen“ Eigenwerte (betragsmäßig kleine Eigen-

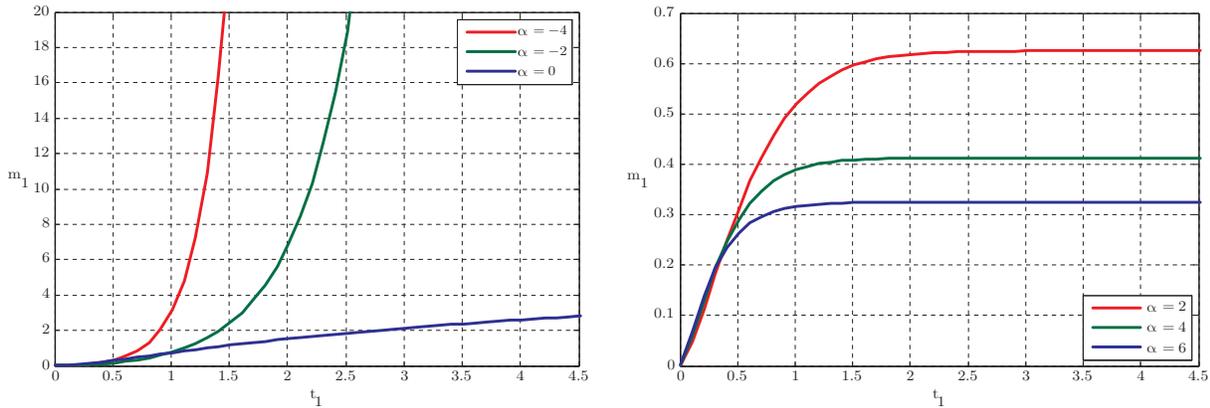


Abbildung 2.3: Steuerbarkeitsmaß nach BENNINGER für unterschiedliche Wahl von t_1

werte) stärker gewichtet. Ist t_1 groß, so haben auch die langsamen Eigenwerte „Zeit, ihren Beitrag zur Überführung des Anfangszustandes“ zu leisten. Daraus lässt sich folgendes erkennen: Wählt man t_1 zu klein, geht die Abhängigkeit der Eigendynamik des Systems verloren. Wählt man die Zeit t_1 hingegen zu groß, so führt dies zu einer Überbewertung der Dynamik und der Einfluss der Eingangsmatrix verschwindet.

BENNINGER schlägt als mögliche Wahl für t_1 die dominante Zeitkonstante des betrachteten Systems vor. Soll das Maß für den Reglerentwurf verwendet werden, erachtet er es als zweckmäßig, t_1 gleich der gewünschten Einschwingzeit zu wählen.

Eine weitere Möglichkeit zur Wahl der Endzeit t_1 wird in [44] vorgestellt. Die Idee dieses Ansatzes ist, t_1 so zu wählen, dass der Einfluss der Systemmatrix \mathbf{A} und der Eingangsmatrix \mathbf{B} auf die Berechnung der GRAMschen Matrix \mathbf{Q}_S gleich groß ist. Daraus folgt

$$\|e^{-\mathbf{A}t_1}\| \stackrel{!}{=} \|\mathbf{B}\|. \quad (2.16)$$

In [44] wurde die Norm $\|\mathbf{M}\| = \max_{i,j} |m_{ij}|$ auf Grund ihrer einfachen Berechenbarkeit vorgeschlagen. Wählt man diese Norm, kann Gleichung (2.16) nicht immer erfüllt werden. Wird die Gleichung jedoch folgendermaßen verändert

$$\|e^{-\mathbf{A}t_1}\| \stackrel{!}{=} \|\mathbf{B}\| + \gamma \quad \gamma \geq 0,$$

ist sie für geeignet gewählte Werte γ in Abhängigkeit der Eigenwerte von \mathbf{A} lösbar. Auf Basis dieser Überlegungen ergibt sich für die Wahl der Endzeit t_1 (vgl. [44]):

$$t_{opt} = \begin{cases} -\frac{\ln(\max_{i,j} |b_{i,j}| + 1)}{\lambda_{\min}} & \text{für } \lambda_{\min} < 0 \\ \max_{i,j} |b_{i,j}| + 1 & \text{für } \lambda_{\min} = 0 \\ -\frac{\ln 0.01}{\lambda_{\min}} & \text{für } \lambda_{\min} > 0. \end{cases}$$

Neben der „gleichmäßigen“ Bewertung der Dynamik der Systemmatrix \mathbf{A} und des Einflusses des Eingangs hat diese Wahl den Vorteil, dass sie i.A. zu einer hinreichend kleinen Endzeit t_1

führt, sodass bei der numerischen Berechnung der GRAMschen Matrix keine Probleme auftreten. Das aus [44] entnommene Beispiel mit den Daten

$$\mathbf{A} = \begin{bmatrix} -49 & 24 \\ -64 & 31 \end{bmatrix}$$

$$\mathbf{b} = \begin{bmatrix} 30 \\ 10 \end{bmatrix}$$

$$t_1 = [0,1 \ 1,5]$$

$$t_{opt} = 0,202$$

zeigt die numerischen Probleme, die bei einer „schlechten“ Wahl von t_1 auftreten können (siehe Abbildung 2.4). Für $t_1 > 1$ liefert die numerische Berechnung der Maßzahlen keine brauchbaren Ergebnisse mehr.

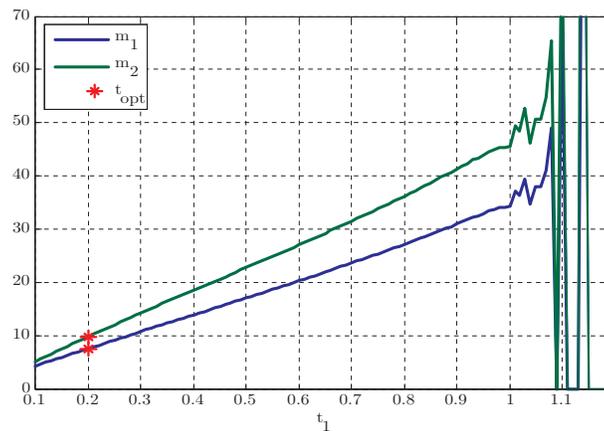


Abbildung 2.4: Maßzahlen für unterschiedliche Werte der Endzeit t_1

2.3.2.2 Berechnung des Steuerbarkeitsmaßes

Für die Berechnung der GRAMschen Matrix nutzt BENNINGER die durch die Zustandstransformation $\mathbf{x} = \mathbf{X}_R \mathbf{z}$ erhaltene Modalform. Die GRAMsche Matrix $\hat{\mathbf{Q}}_S$ eines Systems in Modalform kann analytisch angegeben werden und wird mittels der Beziehung

$$\mathbf{Q}_S = \mathbf{X}_R \hat{\mathbf{Q}}_S \mathbf{X}_R^{-1}$$

rücktransformiert, um die GRAMsche Matrix \mathbf{Q}_S des ursprünglichen Systems zu ermitteln. Da die Elemente der Matrix $\hat{\mathbf{Q}}_S$ i.A. komplex sein können, muss in diesem Fall eine zusätzliche Transformation durchgeführt werden, um eine reelle Darstellung der GRAMschen Matrix zu erhalten (siehe [1]). Der Nachteil bei der Berechnung von \mathbf{Q}_S über die Modalform besteht jedoch darin, dass die Modalform nur existiert, wenn die Systemmatrix \mathbf{A} n voneinander linear unabhängige Eigenvektoren besitzt und kann daher zur Berechnung nicht immer herangezogen werden.

Deshalb wurde in [44] eine andere Methode zur Berechnung der GRAMschen Matrix \mathbf{Q}_S implementiert und getestet. Die zugrunde liegende Idee dieser Methode ist die Darstellung der für

die Berechnung von \mathbf{Q}_S benötigten Transitionsmatrix $\Phi(t) = e^{\mathbf{A}t}$ als TAYLOR-Reihe. Dadurch ergibt sich für die GRAMsche Matrix der Ausdruck

$$\mathbf{Q}_S(t_1) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \left((-1)^{i+j} \frac{1}{i!j!} \mathbf{A}^i \mathbf{B} \mathbf{B}^T (\mathbf{A}^T)^j \frac{t_1^{i+j+1}}{i+j+1} \right),$$

der iterativ berechnet werden kann. Untersuchungen haben ergeben, dass 15 Iterationen i.A. ausreichen, um die ersten drei Kommastellen exakt zu ermitteln. Ferner ergab sich, dass je größer die Endzeit gewählt wird, desto mehr Iterationsschritte benötigt werden, um eine gleich bleibende Genauigkeit zu erzielen [44]. Der Vorteil dieser Methode ist, dass keinerlei Anforderungen an die Systemmatrix \mathbf{A} gestellt werden.

Die für die Berechnung des Steuerbarkeitsmaßes notwendige Inversion der GRAMschen Matrix bereitet vor allem bei Eigenwerten der Matrix \mathbf{A} mit unterschiedlichen Vorzeichen und für große Werte t_1 Schwierigkeiten, da die Größenordnungen der Elemente in \mathbf{Q}_S in diesem Fall sehr unterschiedlich sind. Über eine RICCATI-Gleichung kann die Inverse jedoch direkt berechnet werden: Die GRAMsche Matrix $\mathbf{Q}_S(t)$ erfüllt nämlich die LYAPUNOV-Differentialgleichung

$$\frac{d\mathbf{Q}_S}{dt} = -\mathbf{A}\mathbf{Q}_S - \mathbf{Q}_S\mathbf{A}^T + \mathbf{B}\mathbf{B}^T. \quad (2.17)$$

Berechnet man nun die Ableitung von

$$\mathbf{Q}_S(t)\mathbf{Q}_S^{-1}(t) = \mathbf{E}$$

und setzt Gleichung (2.17) ein, so erfüllt die Matrix $\mathbf{Q}_S^{-1}(t)$ die RICCATI-Differentialgleichung

$$\frac{d\mathbf{Q}_S^{-1}}{dt} = \mathbf{Q}_S^{-1}\mathbf{A} + \mathbf{A}^T\mathbf{Q}_S^{-1} - \mathbf{Q}_S^{-1}\mathbf{B}\mathbf{B}^T\mathbf{Q}_S^{-1}.$$

Nimmt man an, dass für hinreichend große t_1 die Inverse GRAMsche Matrix konstant ist, dann gilt

$$\frac{d\mathbf{Q}_S^{-1}}{dt} = \mathbf{0}$$

und die Matrix \mathbf{Q}_S^{-1} kann über die sogenannte algebraische RICCATI-Gleichung ermittelt werden. Für negative Eigenwerte der Matrix \mathbf{A} ergibt sich für das zugehörige Element in der Matrix \mathbf{Q}_S^{-1} der Wert Null. Das lässt sich leicht dadurch erklären, dass ein asymptotisch stabiles System, das sehr lange Zeit hat ($t_1 \rightarrow \infty$), ohne Einwirkung von außen in den Ursprung gelangt. In diesem Fall ist die benötigte Energie somit gleich Null, und das Steuerbarkeitsmaß wäre damit unendlich groß. Somit liefert die Ermittlung des Maßes über die RICCATI-Gleichung im Falle asymptotisch stabiler Eigenwerte keine brauchbaren Ergebnisse. Jedoch kann die RICCATI-Gleichung zur Bewertung der zu positiven Eigenwerten gehörigen Zustandsvariablen herangezogen werden, wenn die direkte Methode der Invertierung der Matrix \mathbf{Q}_S keine korrekten Ergebnisse liefert. Wird das Maß mit Hilfe der algebraischen RICCATI-Gleichung bestimmt, so entspricht dies der Wahl $t_1 = \infty$.

2.4 Vergleich der Maße

Die Maße sollen bezüglich folgender Forderungen miteinander verglichen werden (vgl. [39]):

- Unabhängigkeit von Zustandstransformationen.
Eine Maßzahl soll für jedes äquivalent transformierte System das gleiche Ergebnis liefern.
- Unabhängigkeit von der Wahl der Eigenvektoren.
Da Links- bzw. Rechts-Eigenvektoren einer Matrix nicht eindeutig sind, ist auch die Modalform eines Systems nicht eindeutig. Die Maßzahlen sollen aber unabhängig von der Transformation in die Modalform das gleiche Ergebnis liefern.
- Konsistenz mit der Definition nach KALMAN.
Das bedeutet, dass das Steuerbarkeitsmaß Null wird, wenn das System die Steuerbarkeit verliert.
- Stetigkeit.
Bei nichtlinearen Systemen sollen linearisierte Modelle, deren Arbeitspunkte „direkt neben einander“ liegen, zu ähnlichen Werten der Maße führen. Natürlich sollen bei linearen zeitinvarianten Systemen, deren Matrizen \mathbf{A} bzw. \mathbf{B} stetig von einem Parameter abhängen, auch die Maße stetige Ergebnisse liefern.
- Physikalische Interpretierbarkeit.
Es soll möglich sein anhand der Maßzahlen direkt auf die Steuerbarkeit der realen nicht transformierten Zustandsvariablen zu schließen.

Es hat sich Folgendes gezeigt: Das Maß a nach KONNO ist invariant gegenüber einer orthogonalen Zustandstransformation, da sich durch die Multiplikation mit einer orthogonalen Matrix die Norm des Eingangsvektors nicht ändert. Die Maße κ nach LÜCKEL und MÜLLER und κ_s nach LITZ sind wegen der Normierung (Division durch der Term $\mathbf{x}_{L_i}^H \mathbf{x}_{L_i}$) ebenfalls invariant gegenüber solchen Transformationen. Alle übrigen modalen Maße weisen eine Unabhängigkeit gegenüber regulären Zustandstransformationen auf. Diese Maße enthalten das Produkt aus transformierter Ein- und Ausgangsmatrix $\hat{\mathbf{C}}\hat{\mathbf{B}}$, welches eine Invariante des Systems darstellt. Bei dem Steuerbarkeitsmaß m nach BENNINGER kann aus methodischen Gründen keine Invarianz gegenüber Transformationen gefordert werden, da in diesem Fall eine Untersuchung der Steuerbarkeit der Zustandsvariablen erfolgt.

Das Regelbarkeitsmaß, die Polempfindlichkeitsmaße und die Dominanzmaße sind auf Grund der Produktbildung $\hat{\mathbf{C}}\hat{\mathbf{B}} = \mathbf{C}\mathbf{X}_R\mathbf{X}_L\mathbf{B} = \mathbf{C}\mathbf{B}$, die Steuerbarkeitsmaße κ und κ_s wegen der Normierung mit \mathbf{X}_L invariant gegenüber der Mehrdeutigkeit der Eigenvektoren. Auch die Energiemaße sind unabhängig von der Mehrdeutigkeit der Eigenvektoren.

Die Forderungen nach Konsistenz und Stetigkeit werden nun anhand von zwei einfachen Beispielen untersucht.

Beispiel 2.2: Gegeben seien die Systemdaten

$$\mathbf{A} = \begin{bmatrix} \alpha & 1 \\ 0 & -2 \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \mathbf{C} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Hierbei ist α eine reelle Konstante. Die Eigenwerte lauten $\lambda_1 = \alpha$ und $\lambda_2 = -2$. Durch Ermittlung der Determinante der Steuerbarkeitsmatrix \mathbf{S}_u

$$|\mathbf{S}_u| = \begin{vmatrix} 1 & \alpha + 1 \\ 1 & -2 \end{vmatrix} = -(\alpha + 3)$$

kann festgestellt werden, dass das System (\mathbf{A}, \mathbf{b}) für $\alpha = -3$ nicht steuerbar ist.

Der Parameter α wird nun von $-5,5$ bis $1,5$ variiert. Abbildung 2.5 zeigt die Steuerbarkeitsmaße κ nach LITZ und κ_s nach LÜCKEL und MÜLLER in Abhängigkeit von α .

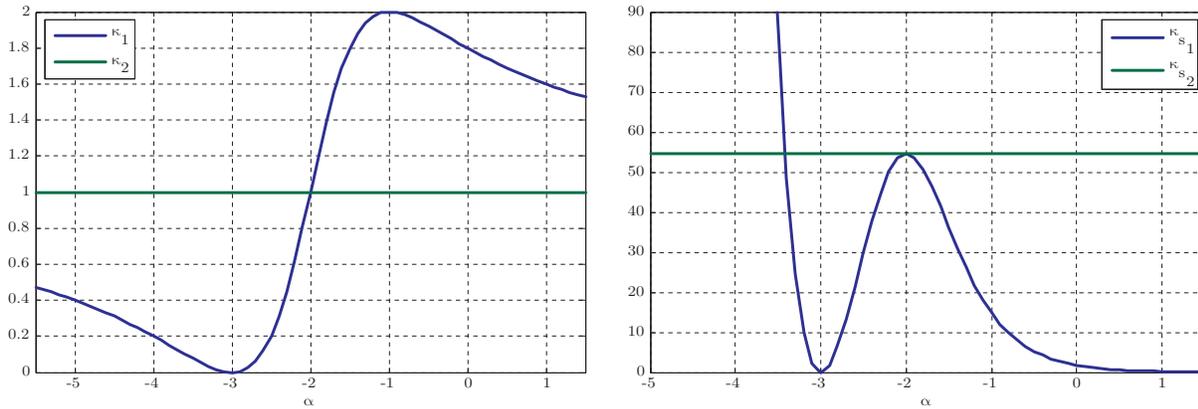


Abbildung 2.5: Steuerbarkeitsmaß κ nach LITZ (links) und κ_s nach LÜCKEL und MÜLLER (rechts)

Beide Maße für den Eigenwert λ_1 sind für $\alpha = -3$ gleich 0, d.h., sie zeigen den Steuerbarkeitsverlust korrekt an und sind somit konsistent zu den KALMANschen Begriffen. Dies gilt in diesem Beispiel auch für alle anderen Maße. Der Term $e^{-2\text{Re}\{\lambda_i\}}$ in der Maßdefinition von κ_s nach MÜLLER und LÜCKEL bewirkt eine Verkopplung der Steuerbarkeit mit der Stabilität des betrachteten Eigenwertes. Dies führt, wie in Abbildung 2.5 zu sehen, zu Verzerrungen des Steuerbarkeitsmaßes. Für $\alpha < -3$ steigt das Maß sehr rasch an und strebt für positive Werte von α gegen 0. Instabile Eigenwerte werden dadurch prinzipiell als schlecht steuerbar beurteilt. Im Allgemeinen ist eine Verkopplung zwischen Steuerbarkeit und Stabilität nicht wünschenswert, da beide Systemeigenschaften unabhängig voneinander sind.

Das System hat an der Stelle $\alpha = -2$ einen doppelten Eigenwert, der, wie in Abbildung 2.6 beispielhaft für das Maß a nach KONNO dargestellt, zu einer Unstetigkeitsstelle führt. In der Nähe dieser Stelle nehmen die Maße sehr hohe Werte an, was fälschlicherweise als „gute Steuerbarkeit“ interpretiert werden kann. Die Unstetigkeit bei mehrfachen Eigenwerten zeigen all die Maße, die auf der Modalform bzw. dem Kriterium von GILBERT fußen, da dieses nur für einfache Eigenwerte definiert ist. Auch die Maße κ und κ_s besitzen für mehrfache Eigenwerte eine Unstetigkeitsstelle bzw. sind für diesen Fall nicht definiert. Diese Unstetigkeitsstelle ist allerdings in Abbildung 2.5 nicht zu sehen, da die Inverse \mathbf{X}_L der singulären Matrix \mathbf{X}_R numerisch berechnet wird. Das so erhaltene \mathbf{X}_L besitzt zwar sehr hohe Werte, die sich jedoch durch die Normierung mit x_{Li} nicht auf das Maß auswirken. Daher ist das Problem der zu hohen Werte in der Umgebung der Unstetigkeitsstelle nicht gegeben.

Das in Abbildung 2.6 dargestellte relative Polempfindlichkeitsmaß nach LITZ $s_{r11}^{\lambda_1}$ zeigt an der Stelle $\alpha = 0$ eine weitere Unstetigkeitsstelle, die aus der Normierung mit dem Eigenwert resultiert. Alle Maße deren Definition eine solche Normierung ausweist, zeigen diese Unstetigkeit.

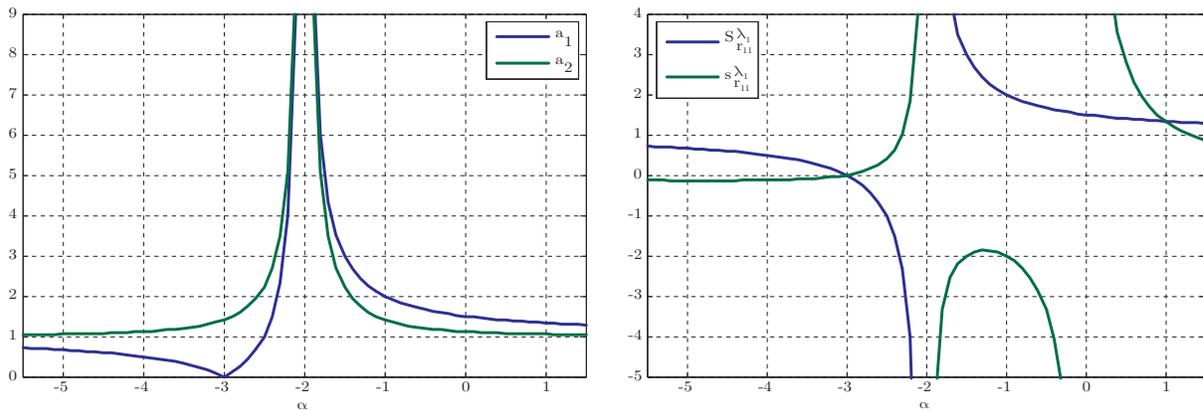


Abbildung 2.6: Steuerbarkeitsmaß a nach KONNO (links) und Polempfindlichkeitsmaße S und s nach LITZ (rechts)

Beispiel 2.3: Gegeben sei das in Modalform vorliegende System

$$\mathbf{A} = \begin{bmatrix} -1 & 0 \\ 0 & -1 + \alpha \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \mathbf{C} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

mit den Eigenwerten $\lambda_1 = -1$, $\lambda_2 = -1 + \alpha$ und einer reellen Konstante α . Für $\alpha = 0$ besitzt das System den doppelten Eigenwert $\lambda_{1,2} = -1$ und ist nicht steuerbar. Beide Zustandsvariablen können über den Stelleingang u beeinflusst werden. Es kann jedoch trotzdem nicht jeder beliebige Zustand erreicht werden, da die beiden entkoppelten Teilsysteme identisch sind und somit das gleiche dynamische Verhalten aufweisen.

Diesen Sachverhalt geben nur die auf der Idee der minimalen Energie basierenden Maße korrekt wieder. Beispielhaft sind in Abbildung 2.7 die Steuer- und Beeinflussbarkeitsmaße m und v nach BENNINGER und RIVOIR dargestellt. Das Steuerbarkeitsmaß m zeigt an der Stelle $\alpha = 0$ den Steuerbarkeitsverlust an. Das Beeinflussbarkeitsmaß v zeigt, dass das System im betrachteten Intervall $\alpha \in [-0,5, 0,5]$ beeinflussbar bleibt.

Da das System in Modalform vorliegt, erhält man konstante, von α unabhängige Eigenvektoren. Daraus resultiert, dass die Modalmaße den Steuerbarkeitsverlust bei $\alpha = 0$ nicht wiedergeben können (vgl. Abbildung 2.7). Dies lässt den Schluss zu, dass die Maße für den Fall mehrfacher Eigenwerte nicht konsistent zu den KALMANschen Begriffen sind. Da die Maße gerade im Falle des Steuerbarkeitsverlustes durch innere Verkopplungen falsch liegen, liegt der Verdacht nahe, dass nicht die Steuerbarkeit bewertet wird, sondern Aussagen über die Beeinflussbarkeit getroffen werden [2]. Für den Fall des Steuerbarkeitsverlustes bei einem doppelten Eigenwert sind somit nur die auf der Idee der minimalen Energie basierenden Maße konsistent. Dies liegt daran, dass für die Berechnung der Maße die GRAMsche Matrix benötigt wird, die für ein nicht steuerbares System singular ist. Daraus folgt, dass die Determinante der Matrix gleich Null ist und die Spur ihrer Inversen theoretisch unendlich groß wird. Dadurch werden die beiden Maße d_1 und d_2 nach KALMAN für nicht steuerbare Systeme zu Null.

Die in diesem Abschnitt untersuchten Eigenschaften der modalen Maße sind in Tabelle 2.1 zusammengefasst.

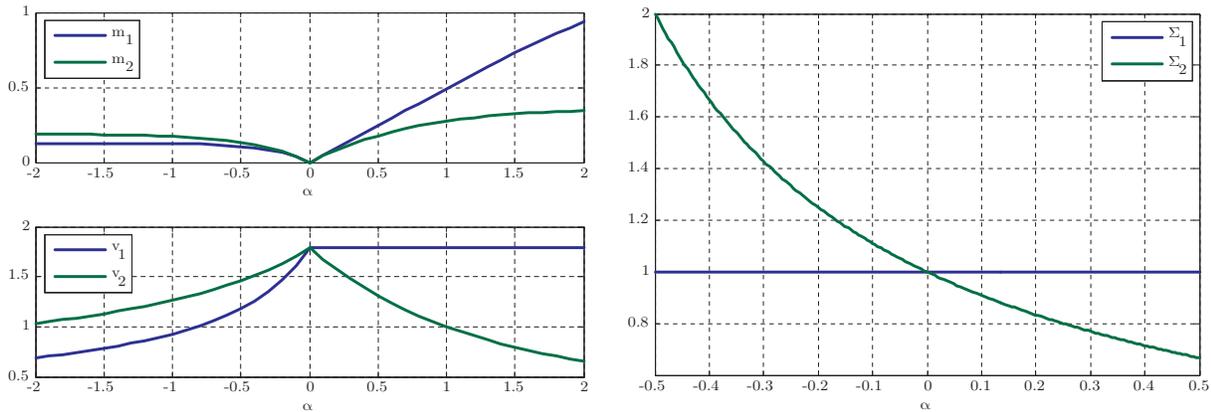


Abbildung 2.7: Das Summenmaß Σ nach LITZ als Beispiel eines nicht konsistenten Maßes (links) und das konsistente Maß m nach BENNINGER und RIVOIR (rechts)

Tabelle 2.1: Vergleich der Modalmaße

	Invariant gegenüber			
	Mehrd. der EV	Koord.trans.	Konsistenz	Stetigkeit
KONNO a (Gl. 2.2)	x	orthogonal	x	x
LÜCKEL, MÜLLER κ_s (Gl. 2.3)	✓	orthogonal	x	x*
LITZ κ (Gl. 2.4)	✓	orthogonal	x	x*
HIPPE ϱ (Gl. 2.5)	✓	regulär	x	x
BENNINGER m (Gl. 2.14)	✓	–	✓	✓**
BENNINGER v (Gl. 2.15)	✓	–	–	✓**
KALMAN:				
Determinante d_1 (Gl. 2.12)	✓	regulär	✓	✓
Spur d_2 (Gl. 2.13)	✓	regulär	✓	✓
LITZ:				
abs. Polempf. S (Gl. 2.6)	✓	regulär	x	x
rel. Polempf. s (Gl. 2.7)	✓	regulär	x	x
Maximalmaß M (Gl. 2.8)	✓	regulär	x	x
Summenmaß Σ (Gl. 2.9)	✓	regulär	x	x

* Besitzt Unstetigkeitsstelle für mehrfache Eigenwerte; durch numerische Fehler bei der Inversion von \mathbf{X}_R erhält man jedoch meist einen stetigen Funktionsverlauf.

** In manchen Fällen Unstetigkeiten durch numerische Fehler bei der Berechnung.

Zum Abschluss dieses Abschnittes wird untersucht, ob es sich bei den betrachteten Maßzahlen auch um ein Maß bzw. eine Norm im mathematischen Sinn handelt. Hierzu folgende Definition:

Definition 1 (Norm): Eine Abbildung $\|\cdot\| : V \rightarrow \mathbf{R}_0^+$ in die nichtnegativen reellen Zahlen heißt Norm auf V , wenn für alle Vektoren $\mathbf{x}, \mathbf{y} \in V$ und alle Skalare λ die folgenden Bedingungen erfüllt sind:

1. $\|\mathbf{x}\| = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$ (Definitheit).

2. $\|\mathbf{x}\| > 0 \Leftrightarrow \mathbf{x} \neq \mathbf{0}$ (Positivität).
3. $\|\lambda\mathbf{x}\| = |\lambda|\|\mathbf{x}\|$ (Homogenität).
4. $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ (Dreiecksungleichung).

Handelt es sich um eine Matrixnorm, so gilt zusätzlich für zwei Matrizen \mathbf{X}, \mathbf{Y} passender Dimension

5. $\|\mathbf{XY}\| \leq \|\mathbf{X}\|\|\mathbf{Y}\|$ (Submultiplikativität).

Für die modalen Steuerbarkeitsmaße wird überprüft, ob das Maß eine Norm des Zeilenvektors $\hat{\mathbf{b}}_i^T$ der transformierten Eingangsmatrix darstellt. Das Maß κ nach LITZ erfüllt die Eigenschaft der Definitheit und der Positivität. Die beiden übrigen Eigenschaften werden von dieser Maßzahl nicht erfüllt. Wählt man jedoch stattdessen als Maß $\sqrt{\kappa}$, so können alle vier Eigenschaften erfüllt werden. Die Erfüllung der Eigenschaft 4 ist unter der Voraussetzung gleicher Links-Eigenvektoren möglich. D.h., wenn V die Menge der Eingangsvektoren *eines Systems* beschreibt, dann ist das Maß $\sqrt{\kappa}$ eine Norm auf V . Analoges gilt auch für das Maß κ_S nach MÜLLER und LÜCKEL. Für die Maßzahl $\sqrt{\kappa_s}$ müssen zur Erfüllung der Dreiecksungleichung sowohl die Links-Eigenvektoren als auch die Eigenwerte übereinstimmen. Das Maß a nach KONNO erfüllt alle vier Eigenschaften.

Bei den Steuer- und Beobachtbarkeitsmaßen ist zu prüfen, ob es sich um eine Matrixnorm handelt, da sich in diesem Fall das Maß auf die Matrix \mathbf{G}_i aus Gleichung (2.2) bezieht. Da für das Regelbarkeitsmaß nach HIPPE und für die Polempfindlichkeitsmaße nach LITZ die einzelnen Elemente der Matrix \mathbf{G}_i betrachtet werden, handelt es sich bei diesen Maßen um keine Normen, da sie die Abbildung $\mathbb{R}^{p \times m} \rightarrow \mathbb{R}^{p \times m}$ beschreiben. Das Maximal- und das Summenmaß M_i und Σ_i nach LITZ erfüllen die Eigenschaften 1-3. Für gleiche Eigenwerte λ_i wird auch die 4. Eigenschaft erfüllt. Die Eigenschaft der Submultiplikativität erfüllen diese Maße nicht. Unter der Voraussetzung, dass das System gleich viele Ein- wie Ausgänge ($m = p$) besitzt, erfüllt das Produkt aus der Anzahl der Eingänge m und dem Maximalmaß M_i auch diese Eigenschaft. Somit stellt mM_i eine Matrixnorm dar.

Die über die minimale Energie definierten Maße nach BENNINGER und RIVOIR sind keine Normen, da sie die GRAMsche Matrix \mathbf{Q}_S auf n Werte abbilden.

2.4.1 Mögliche Erweiterung

Die Untersuchung der Modalmaße hat gezeigt, dass die Verwendung einiger Maße zu falschen Ergebnissen führt, wenn die Eigenwerte der Systemmatrix nahe beieinanderliegen. Besitzt die Matrix \mathbf{A} einen doppelten Eigenwert, dann haben diese Maße eine Unstetigkeitsstelle und der Steuerbarkeitsverlust auf Grund eines doppelten Eigenwertes wird vom Maß nicht angezeigt. In der Umgebung dieser Unstetigkeitsstelle, d.h., wenn der Abstand zwischen den Eigenwerten klein ist, liefern diese Maße fälschlicherweise sehr hohe Werte. Leider kann kein maximaler Abstand angegeben werden, den die Eigenwerte besitzen müssen, damit dieser Effekt nicht auftritt. Durch Multiplikation des Maßes mit dem kleinsten Abstand

$$\delta_\lambda := \min\{|\lambda_i - \lambda_j|\} \quad \text{für } i \neq j \quad i, j = 1, \dots, n$$

zwischen zwei Eigenwerten kann diesem Effekt entgegengewirkt werden. In diesem Abschnitt wird die Auswirkung der Multiplikation mit δ_λ am Beispiel des Maßes a nach KONNO untersucht. Für die sich aus der Multiplikation ergebende Funktion wird die Bezeichnung

$$\tilde{a} := a\delta_\lambda$$

eingeführt.

Das in Beispiel 2.3 untersuchte System verliert für $\alpha = 0$ wegen des doppelten Eigenwertes $\lambda_{1,2} = -1$ die Steuerbarkeit. Dies wird vom Maß a nicht angezeigt, da es für alle α konstant gleich Eins ist. Der Abstand zwischen den Eigenwerten ist für $\alpha = 0$ gleich Null, und somit ist auch \tilde{a} gleich Null (vgl. Abbildung 2.8). Durch die Multiplikation von a mit δ_λ wird für dieses Beispiel der Steuerbarkeitsverlust nun korrekt angezeigt.

Das System aus Beispiel 2.2 verliert für $\alpha = -3$ die Steuerbarkeit und hat für $\alpha = -2$ einen doppelten Eigenwert. Da das Maß a den Steuerbarkeitsverlust für $\alpha = -3$ korrekt wiedergibt (vgl. Abbildung 2.6), wird auch \tilde{a} zu Null. Für den doppelten Eigenwert bei $\alpha = -2$ ist das System steuerbar. Da gilt $\delta_\lambda(\alpha = -2) = 0$, wird auch $\tilde{a}(\alpha = -2)$ zu Null, wodurch fälschlicherweise angezeigt wird, dass das System für $\alpha = -2$ nicht steuerbar ist.

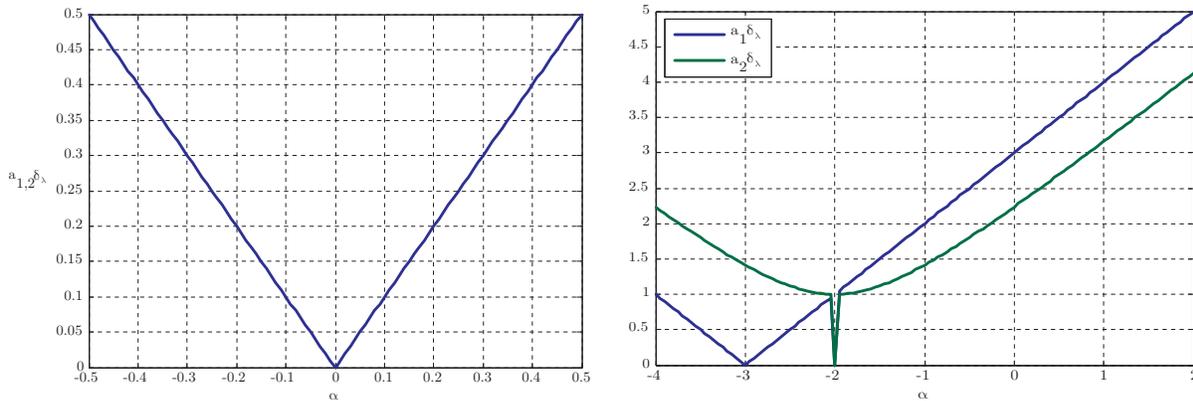


Abbildung 2.8: Funktion $a\delta_\lambda$ für System aus Beispiel 2.2 (rechts) und Beispiel 2.3 (links)

Die Multiplikation von a mit δ_λ hat die zuvor gewünschte Auswirkung: \tilde{a} besitzt immer noch eine Unstetigkeitsstelle für mehrfache Eigenwerte (siehe Abbildung 2.8 (rechts) für $\alpha = -2$). Den fälschlicherweise hohen Werten in der Umgebung der Unstetigkeitsstellen von a wird allerdings entgegengewirkt. Der Nachteil von \tilde{a} ist, dass im Falle eines doppelten Eigenwertes $\delta_\lambda = 0$ gilt. Dadurch zeigt \tilde{a} immer an, dass das System nicht steuerbar ist, unabhängig davon, ob der doppelte Eigenwert zu einem Steuerbarkeitsverlust führt oder nicht.

2.4.2 Diskussion der Ergebnisse

Die in diesem Kapitel vorgestellten Maße lassen sich in zwei Klassen einteilen: *Modalmaß* und *Energiemaß*. Die Klasse der Modalmaße zeichnet sich durch einfache Berechenbarkeit aus, besitzt aber auch Schwächen: Sie beschränkt sich auf diagonalisierbare Systemmatrizen \mathbf{A} . Dadurch ergeben sich Unstetigkeitsstellen in den Maßen. Bei einigen Maßdefinitionen ergeben sich weitere Unstetigkeiten durch die Normierung mit dem Eigenwert. Es hat sich gezeigt, dass

die Maße in der Nähe dieser Unstetigkeitsstellen i.A. sehr hohe Werte annehmen, was fälschlicherweise als „gute Steuerbarkeit“ interpretiert werden kann. Da diese Definitionen nur die Eingangsmatrix \mathbf{B} berücksichtigen, zeigen die Maße einen Steuerbarkeitsverlust auf Grund innerer Verkopplungen, hervorgerufen durch mehrfache Eigenwerte, nicht an. Durch die Betrachtung im Modalraum ist i.A. eine Zuordnung zu den ursprünglichen (physikalischen) Zustandsvariablen nicht möglich.

Betrachtet man die Klasse der Energiemaße, erkennt man, dass die Stabilität des Systems einen Einfluss auf das Steuerbarkeitsmaß hat. Liegt z.B. ein asymptotisch stabiles System vor, so wird bei konstantem Endzeitpunkt t_1 das Maß größer, je weiter die Eigenwerte des Systems von der imaginären Achse entfernt sind. Der Grund hierfür ist, dass je „stabiler“ ein Eigenwert, desto schneller wird der Anfangszustand \mathbf{x}_0 in den Ursprung überführt. Somit ist die über den Eingang \mathbf{u} zugeführte Energie klein. Dieser Effekt verstärkt sich, je größer t_1 gewählt wird. Die Verkopplung zwischen Steuerbarkeit und Stabilität kommt daher, dass man in dieser Definitionsklasse die minimale Energie betrachtet, die benötigt wird, um einen Anfangszustand *in den Ursprung* zu überführen. Bei der Definition der Steuerbarkeit wird jedoch die Überführbarkeit eines Anfangszustandes in jeden beliebigen Endzustand gefordert. Je nach Lage von Anfangs- und Endzustand kann auch ein instabiler Eigenwert dazu beitragen, dass die benötigte minimale Energie klein ist. Das wird von den Energiemaßen nicht berücksichtigt. Diese Maße bewerten genau genommen nicht die Steuerbarkeit, sondern die „Steuerbarkeit in den Ursprung“. Die Berechnung der Maße ist im Vergleich zu den Modalmaßen aufwendiger, da die Inverse der GRAMschen Matrix berechnet werden muss. In [1] wird für die Berechnung die Modalform benutzt, da für diese Systeme die GRAMsche Matrix analytisch angegeben werden kann. Damit ergibt sich aber der Nachteil, dass die Berechnung nur für diagonalisierbare Systemmatrizen möglich ist. Im Unterschied zu den Modalmaßen ist die Definition an sich nicht auf diese Systemklasse beschränkt, sondern diese Einschränkung ergibt sich erst durch die Berechnung. Dieses Problem kann durch die in Abschnitt 2.3.2.2 vorgestellte Berechnung über den Reihenansatz umgangen werden.

Kapitel 3

Distanzmaß

Da das in der Einleitung angeführte Theorem 1 nach LEE und MARKUS für ein tieferes Verständnis des Steuerbarkeitsbegriffes von großer Bedeutung ist, wird dieses hier erneut angeführt und erläutert.

Theorem 4 (LEE und MARKUS [28]): Betrachtet man ein steuerbares System (\mathbf{A}, \mathbf{B}) , so existiert *eine* positive Konstante ϵ_s derart, dass *jedes* System $(\hat{\mathbf{A}}, \hat{\mathbf{B}})$ mit der Eigenschaft

$$\|\mathbf{A} - \hat{\mathbf{A}}\| < \epsilon_s \quad \text{und} \quad \|\mathbf{B} - \hat{\mathbf{B}}\| < \epsilon_s \quad (3.1a)$$

ebenfalls steuerbar ist. Umgekehrt, falls ein System $(\tilde{\mathbf{A}}, \tilde{\mathbf{B}})$ nicht steuerbar ist, so existiert zu *jeder* positiven Konstante ϵ ein steuerbares System (\mathbf{A}, \mathbf{B}) mit der Eigenschaft

$$\|\mathbf{A} - \tilde{\mathbf{A}}\| < \epsilon \quad \text{und} \quad \|\mathbf{B} - \tilde{\mathbf{B}}\| < \epsilon. \quad (3.1b)$$

Relation (3.1a) sagt aus, dass um jedes steuerbare System eine Umgebung existiert, in welcher sich nur steuerbare Systeme befinden. In Relation (3.1b) kann der Parameter ϵ beliebig klein gewählt werden. Somit besagt diese Ungleichung, dass sich beliebig nahe bei einem nicht steuerbaren System ein steuerbares befindet (vgl. Abbildung 3.1).

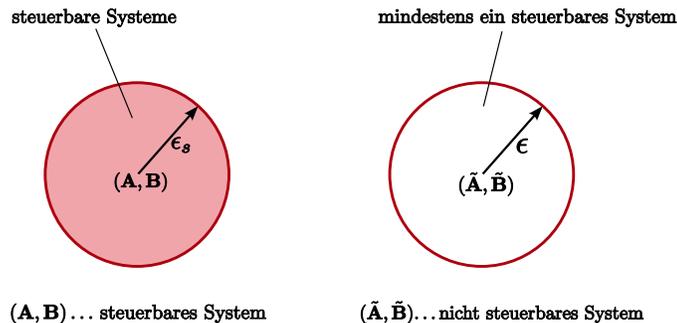


Abbildung 3.1: Theorem nach LEE und MARKUS

Damit ist die Menge der steuerbaren Systeme offen und dicht im metrischen Raum aller LZI-Systeme. In diesem Raum befinden sich nur an singulären Punkten nicht steuerbare Systeme. Es wird nun folgende Größe eingeführt:

$$\mu(\mathbf{A}, \mathbf{B}) := \min_{\delta \mathbf{A} \in \mathbb{C}^{n \times n}, \delta \mathbf{B} \in \mathbb{C}^{n \times m}} \{ \|\delta \mathbf{A}, \delta \mathbf{B}\| : (\mathbf{A} + \delta \mathbf{A}, \mathbf{B} + \delta \mathbf{B}) \text{ nicht steuerbar} \}. \quad (3.2)$$

Mit der Spektralnorm einer (rechteckigen) Matrix ¹

$$\|\mathbf{M}\| := \sqrt{\lambda_{\max}(\mathbf{M}^H \mathbf{M})} = \sigma_{\max}(\mathbf{M}),$$

die dem größten Singulärwert der Matrix \mathbf{M} entspricht. Aus dem Theorem nach LEE und MARKUS ist erkennbar, dass die Größe μ dem kleinsten Abstand zu einem nicht steuerbaren System entspricht, weshalb sie als *Distanzmaß* bezeichnet wird. Ist dieser kleinste Abstand zu einem nicht steuerbaren System klein, wird das System als „schlecht“ steuerbar bezeichnet.

Diese Größe kann jedoch auch anders interpretiert werden: In Relation (3.2) wird nach einer Störung $[\delta\mathbf{A}, \delta\mathbf{B}]$ mit der kleinstmöglichen Norm gesucht, die das System (\mathbf{A}, \mathbf{B}) nicht steuerbar macht. Ist das Maß klein, heißt das, dass eine kleine Störung ausreicht, um das System $(\mathbf{A} + \delta\mathbf{A}, \mathbf{B} + \delta\mathbf{B})$ nicht steuerbar zu machen. In diesem Fall wird das System (\mathbf{A}, \mathbf{B}) als „schlecht“ steuerbar bezeichnet.

PAIGE führt in [38] die Größe μ für eine numerisch zuverlässige Überprüfung der Steuerbarkeit ein, da, wie in der Einleitung gezeigt wurde, die numerische Überprüfung der Steuerbarkeit mit Hilfe der Kriterien nach KALMAN oder HAUTUS in manchen Fällen zu fehlerhaften Ergebnissen führt. Ist $\mu(\mathbf{A}, \mathbf{B})$ größer als ein bestimmter Wert μ_0 , der sich aus Unsicherheiten der Systemmatrizen und aus einer Schranke für die zu erwartenden numerischen Fehler bei der Berechnung von μ zusammensetzt, dann ist das System (\mathbf{A}, \mathbf{B}) mit Sicherheit steuerbar. Andernfalls kann keine zuverlässige Aussage über die Steuerbarkeit des Systems gemacht werden.

EISING hat in [9, 10] gezeigt, dass Relation (3.2) zu der Beziehung

$$\mu(\mathbf{A}, \mathbf{B}) = \inf_{s \in \mathbb{C}} \{\sigma_{\min}([s\mathbf{E} - \mathbf{A}, \mathbf{B}])\} =: \inf_{s \in \mathbb{C}} \{\sigma(s)\} \quad (3.3)$$

äquivalent ist. Hierbei ist s eine beliebige komplexe Zahl und $\sigma_{\min}(\mathbf{M})$ der kleinste Singulärwert der Matrix \mathbf{M} . Diese Gleichung beschreibt ein Optimierungsproblem mit dem freien Parameter s . Somit erhält man unabhängig von der Systemordnung ein Problem mit zwei reellen Optimierungsparametern: dem Real- und dem Imaginärteil der komplexen Variablen s . Im Vergleich dazu kann im Optimierungsproblem nach (3.2) an den Matrizen $\delta\mathbf{A}$ und $\delta\mathbf{B}$ „gedreht“ werden, woraus sich $n(n + m)$ Optimierungsparameter ergeben.

Die Zusammenhänge zwischen (3.2) und (3.3) können leicht gezeigt werden: Angenommen das System $(\mathbf{A} + \delta\mathbf{A}, \mathbf{B} + \delta\mathbf{B})$ ist nicht steuerbar. Dann gilt für mindestens ein $s \in \mathbb{C}$

$$\text{Rang}([s\mathbf{E} - (\mathbf{A} + \delta\mathbf{A}), \mathbf{B} + \delta\mathbf{B}]) < n.$$

Da die kleinste Störung, die einen Rangverlust der Matrix $[s\mathbf{E} - \mathbf{A}, \mathbf{B}]$ hervorrufen kann, gleich $\sigma_{\min}([s\mathbf{E} - \mathbf{A}, \mathbf{B}])$ ist, gilt

$$\sigma_{\min}([s\mathbf{E} - \mathbf{A}, \mathbf{B}]) \leq \|[\delta\mathbf{A}, \delta\mathbf{B}]\|.$$

Das Gleichheitszeichen gilt für

$$[\delta\mathbf{A}, \delta\mathbf{B}] = -\sigma_{\min} \mathbf{u}_n \mathbf{v}_n^H$$

mit den zum Singulärwert σ_{\min} gehörigen normierten Links- und Rechtssingulärvektoren \mathbf{u}_n und \mathbf{v}_n .

¹Hierbei symbolisiert $\lambda_{\max}(\mathbf{M})$ den größten Eigenwert der Matrix \mathbf{M} .

Gleichung (3.3) besagt, dass μ das Minimum des durch $\sigma_{\min}([s\mathbf{E} - \mathbf{A}, \mathbf{B}])$ über der komplexen Ebene aufgespannten „Gebirges“ ist. Die Visualisierung dieses „Gebirges“ kann mit Hilfe des sogenannten Pseudospektrums erfolgen (siehe hierfür Anhang A.1).

Eine einfache obere Schranke für das Distanzmaß ist durch die Norm der Matrix \mathbf{A}

$$\mu(\mathbf{A}, \mathbf{B}) \leq \|\mathbf{A}\|$$

gegeben [16].

Der Bereich, in dem sich das Minimum befindet, kann wie folgt eingeschränkt werden: Es sei

$$f(s) := \mathbf{v}_n^H(s) \begin{bmatrix} \mathbf{u}_n(s) \\ 0 \end{bmatrix} \quad (3.4)$$

mit $\mathbf{u}_n(s)$ und $\mathbf{v}_n(s)$ den normalisierten n -ten Spalten der Matrizen \mathbf{U} und \mathbf{V} aus der Singulärwertzerlegung

$$[s\mathbf{E} - \mathbf{A}, \mathbf{B}] = \mathbf{U}\Sigma\mathbf{V}^H.$$

Die *kritischen* Punkte von $\sigma(s)$ genügen der Beziehung

$$s = \mathbf{u}_n(s)^H \mathbf{A} \mathbf{u}_n(s). \quad (3.5)$$

Diese folgt aus

$$\sigma(s)f(s) = \mathbf{u}_n(s)^H (s\mathbf{E} - \mathbf{A}) \mathbf{u}_n(s),$$

da

$$[s\mathbf{E} - \mathbf{A}, \mathbf{B}]^H \mathbf{u}_n(s) = \sigma(s) \mathbf{v}_n(s)$$

gilt. Da alle kritischen Punkte $s = x + jy$ Gleichung (3.5) erfüllen, liegen alle Minima im Wertebereich² von \mathbf{A}

$$F(\mathbf{A}) = \{\mathbf{x}^H \mathbf{A} \mathbf{x} : \|\mathbf{x}\| = 1, \mathbf{x} \in \mathbb{C}\}.$$

Damit und durch Aufspalten der Matrix \mathbf{A} in einen symmetrischen und einen schiefsymmetrischen Anteil kann der Bereich, in dem das Minimum liegen, auf das Rechteck³

$$\begin{aligned} \lambda_{\min} \left(\frac{\mathbf{A} + \mathbf{A}^T}{2} \right) &\leq x \leq \lambda_{\max} \left(\frac{\mathbf{A} + \mathbf{A}^T}{2} \right) \\ \lambda_{\min} \left(\frac{\mathbf{A} - \mathbf{A}^T}{2j} \right) &\leq y \leq \lambda_{\max} \left(\frac{\mathbf{A} - \mathbf{A}^T}{2j} \right) \end{aligned} \quad (3.6)$$

beschränkt werden [7]. Da $\sigma_{\min}([s\mathbf{E} - \mathbf{A}, \mathbf{B}]) = \sigma_{\min}([\bar{s}\mathbf{E} - \mathbf{A}, \mathbf{B}])$ gilt, gibt es nur reelle oder konjugiert komplexe Minima. Deshalb muss nur in einer der beiden Hälften des oben angegebenen Rechtecks nach dem globalen Minimum gesucht werden. Die Schwierigkeit der Minimumsuche ergibt sich dadurch, dass Gleichung (3.3) ein nichtlineares, nicht glattes, nicht konvexes Optimierungsproblem beschreibt.

Zur Bestimmung von μ nach Gleichung (3.2) wird nach der kleinsten Norm einer Störung $[\delta\mathbf{A}, \delta\mathbf{B}]$ gesucht, die das System nicht steuerbar macht. Hierbei werden zwei Fälle unterschieden:

²Zur Definition und zu den Eigenschaften des Wertebereichs einer Matrix siehe [46].

³Da hier nur Systeme mit reellen Daten betrachtet werden, wird die transponierte Matrix benutzt. Andernfalls muss diese durch die konjugiert transponierte Matrix ersetzt werden.

- Die Matrizen $\delta\mathbf{A}$ und $\delta\mathbf{B}$ können komplexe Elemente besitzen. Das sich dadurch ergebende μ wird im Weiteren als μ_C bezeichnet.
- Die Matrizen $\delta\mathbf{A}$ und $\delta\mathbf{B}$ beinhalten nur reelle Elemente, da im Allgemeinen reelle Störungen von praktischem Interesse sind. Das Optimierungsproblem zur Bestimmung des als μ_R bezeichneten Maßes entspricht einer Minimumsuche entlang der reellen Achse des Pseudospektrums.

Wird in weiterer Folge vom Distanzmaß μ (ohne Indizierung) gesprochen, dann ist sowohl μ_C als auch μ_R gemeint.

In den folgenden Abschnitten wird gezeigt, wie das Maß μ_C bzw. μ_R ermittelt werden kann. Für bestimmte Systeme 1. und 2. Ordnung kann das Distanzmaß analytisch angegeben werden (siehe Abschnitt 3.1). Die analytische Berechnung wird jedoch mit zunehmender Ordnung schwieriger bzw. unmöglich. Das Maß μ wird dann numerisch ermittelt. In Abschnitt 3.2.1 werden Algorithmen vorgestellt, die eine obere und untere Schranke für das zu ermittelnde Minimum μ_C liefern. In Abschnitt 3.2.2 werden Verfahren zur Bestimmung von Schranken für μ_R vorgestellt.

3.1 Analytische Lösung

Die analytische Lösung liefert einen tieferen Einblick und ein besseres Verständnis für das Distanzmaß. Zuerst werden Systeme 1. Ordnung und in Modalform vorliegende Systeme 2. Ordnung betrachtet, da für diese Systeme das Distanzmaß in einigen Fällen durch Lösen des Optimierungsproblems (3.3) nach EISING analytisch ermittelt werden kann. Mit Hilfe dieser analytischen Lösungen und weiterer Abschätzungen können dann auch für allgemeine Systeme höherer Ordnung Aussagen über das Verhalten des Distanzmaßes gemacht werden.

3.1.1 Systeme 1. Ordnung

Für das System 1. Ordnung

$$\frac{dx}{dt} = ax + bu$$

wird nun das Distanzmaß

$$\mu = \inf_{s \in \mathbb{C}} \{\sigma_{\min}([s - a, b])\} = \inf_{s \in \mathbb{C}} \{\sigma(s)\} = \sigma(s^*) \quad (3.7)$$

mit $s = x + jy$ analytisch berechnet.

Mit der Abkürzung $\mathbf{M} := [s - a, b]$ ergibt sich für den Singulärwert σ

$$\sigma(x, y) = \sqrt{\lambda_{\max}(\mathbf{M}\mathbf{M}^H)} = \sqrt{(x - a)^2 + y^2 + b^2}.$$

Das Rechteck (3.6) beschreibt den Bereich in der komplexen Ebene in dem sich das Minimum befindet. Für Systeme 1. Ordnung entartet dieses Rechteck zu einer Geraden entlang der reellen

Achse. Somit gilt für die Lage des Minimums $s^* = x^*$. Um dies zu berechnen, wird σ für $y = 0$ nach x abgeleitet und die Ableitung Null gesetzt. Dadurch ergibt sich

$$\frac{d\sigma}{dx} \stackrel{!}{=} 0 \quad \leadsto \quad x^* = a.$$

Da die 2. Ableitung von σ an der Stelle $x = a$ positiv ist, handelt es sich bei der gefundenen Lösung tatsächlich um ein Minimum. Setzt man dieses Ergebnis in Gleichung (3.7) ein, so erhält man für das Distanzmaß

$$\mu = \sigma([a - a, b]) = |b|.$$

Aus diesem Ergebnis ist ersichtlich, dass der Eigenwert a des Systems die Lage des Minimums s^* bestimmt, während die Größe des Distanzmaßes μ nur vom Betrag von b abhängig ist. Da $y = 0$ gilt, ist s^* reell und somit $\mu_C = \mu_R$.

Für das System 1. Ordnung wird zum Vergleich das Steuerbarkeitsmaß κ nach LITZ berechnet. Da für die Eigenvektoren $x_R = x_L = 1$ gilt, vereinfacht sich die Berechnung von κ zu

$$\kappa = \frac{x_L^H b b^H x_L}{x_L^H x_L} = b^2.$$

Daher gilt zwischen κ und μ der einfache Zusammenhang $\mu^2 = \kappa$.

Zusammenfassung (Systeme 1. Ordnung):

1. $\mu_C = \mu_R$.
2. a bestimmt die Lage des Minimums s^* .
3. b bestimmt das Steuerbarkeitsmaß μ .
4. $\mu^2 = \kappa$.

3.1.2 Systeme 2. Ordnung

In diesem Abschnitt wird die analytische Ermittlung des Distanzmaßes μ für Systeme 2. Ordnung angegeben. Der Einfachheit halber wird das Distanzmaß μ für das in Modalform vorliegende System

$$\frac{dx}{dt} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \mathbf{x} + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} u \quad (3.8)$$

ermittelt. Mit der Abkürzung $\mathbf{M} := [s\mathbf{E} - \mathbf{A}, \mathbf{b}]$ berechnen sich die Singulärwerte der Matrix \mathbf{M} mit $s = x + jy$ gemäß

$$\sigma(x, y) = \sqrt{\lambda(\mathbf{M}\mathbf{M}^H)}.$$

3.1.2.1 Reelle Eigenwerte

Besitzt die Matrix \mathbf{A} reelle Eigenwerte, entartet das Rechteck (3.6) in dem sich das Minimum befindet zu einer Geraden auf der reellen Achse. Damit gilt für die Lage des Minimums $s^* = x^* \in \mathbb{R}$ und somit $\mu_C = \mu_R$. Über die analytische Lösung kann folgendes festgehalten werden: Es gilt $\lambda_1, \lambda_2, b_1, b_2 \in \mathbb{R}$ und man erhält

$$\mathbf{MM}^H = \begin{bmatrix} \alpha_1 & b_1 b_2 \\ b_1 b_2 & \alpha_2 \end{bmatrix} \quad (3.9)$$

mit

$$\alpha_1 := (x - \lambda_1)^2 + y^2 + b_1^2 \quad \text{und} \quad \alpha_2 := (x - \lambda_2)^2 + y^2 + b_2^2.$$

Für die Ableitung von $\sigma(x, y = 0)$ nach x ergibt sich ein Polynom 4. Grades, dessen Nullstellen nicht analytisch angegeben werden können. Für den Spezialfall, dass die Systemmatrix \mathbf{A} zwei gleiche Eigenwerte $\lambda_1 = \lambda_2 =: \lambda$ besitzt, vereinfacht sich die Berechnung. Für das globale Minimum gilt in diesem Fall $s^* = \lambda$ mit $\mu(s^*) = 0$. Daraus folgt, dass das Distanzmaß für Systeme 2. Ordnung in Diagonalform den Steuerbarkeitsverlust auf Grund doppelter Eigenwerte korrekt wiedergibt.

Leider kann schon für einfache Systeme 2. Ordnung keine analytische Lösung für das Distanzmaß μ ermittelt werden. Um das Distanzmaß trotzdem analysieren zu können, wird nun eine asymptotische Lösung für das Maß ermittelt. Hierfür wird wieder das System in Modalform (3.8) betrachtet. Der Einfachheit halber soll vorerst $b_1 = b_2 =: b$ gelten. Durch diese vereinfachende Annahme kann eine Lösung für

$$\frac{\partial \sigma}{\partial x} = 0$$

und somit ein lokales Extremum angegeben werden mit

$$x_1^* = \frac{1}{2}(\lambda_1 + \lambda_2).$$

Berechnet man für dieses Extremum den Wert für das Distanzmaß μ , so ergibt sich

$$\mu(x_1^*) = \frac{1}{2}\delta \quad (3.10)$$

mit $\delta = |\lambda_1 - \lambda_2|$, dem Abstand zwischen den beiden Eigenwerten. Diese Gleichung beschreibt eine Asymptote in Abhängigkeit von δ . Eine weitere Asymptote erhält man für die Lösung $x_2^* = \lambda_{1,2}$. Strebt δ gegen Unendlich, dann gilt

$$\lim_{\delta \rightarrow \infty} \left. \frac{\partial \sigma}{\partial x} \right|_{x=x_2^*} = 0.$$

D.h., wird der Abstand zwischen den Eigenwerten der Systemmatrix \mathbf{A} größer, strebt die Lösung x_2^* einem Extremum entgegen. Für den Wert des Distanzmaßes gilt in diesem Fall

$$\lim_{\delta \rightarrow \infty} \mu(x_2^*) = b, \quad (3.11)$$

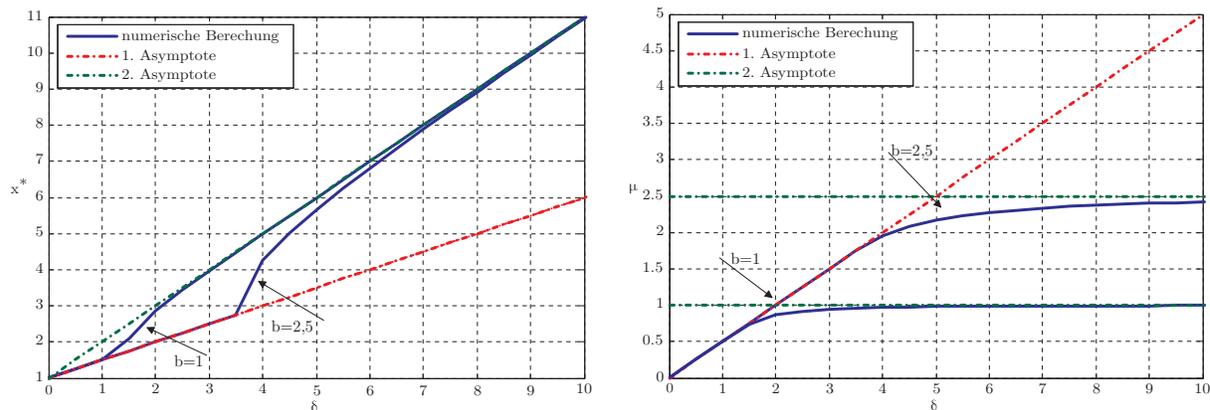


Abbildung 3.2: Numerische und asymptotische Lösung für das Minimum x^* (links) und das Distanzmaß μ (rechts)

welcher die zweite Asymptote beschreibt. Der Schnittpunkt der beiden Asymptoten (3.10) und (3.11) ergibt sich zu $\delta = 2b$.

Abbildung 3.2 zeigt die numerisch bestimmten Werte für die Lage des Minimums x^* , das Distanzmaß μ sowie die ermittelten Asymptoten, aufgetragen über dem Abstand δ zwischen den Eigenwerten der Systemmatrix A . Es ist ersichtlich, dass die asymptotische Darstellung mit größer werdenden Werten für b schlechter wird. Trotzdem liefert diese Darstellung wichtige Einblicke in das Verhalten des Distanzmaßes μ bzw. in das Steuerbarkeitsverhalten des Systems. Ist der Abstand δ hinreichend groß, dann ist der Einfluss der Systemmatrix vernachlässigbar klein, und nur noch der Eingangsvektor \mathbf{b} ist ausschlaggebend für das Distanzmaß μ . Für hinreichend kleine Werte von b liegt das Minimum in der Mitte der beiden Eigenwerte von A , und für den Wert des Distanzmaßes μ ist nur der Abstand δ entscheidend.

Für den Fall $b_1 \neq b_2$ ist die Lösung $\mu(x_1^*) = \frac{1}{2}\delta$ kein Extremum des Distanzmaßes μ . Jedoch liefert die zuvor ermittelte Asymptote aus Gleichung (3.10) auch in diesem Fall eine brauchbare Abschätzung. Für $x_2^* = \lambda_{1,2}$ gilt analog zum Fall $b_1 = b_2$, dass diese Lösung für $\delta \rightarrow \infty$ einem Extremum mit dem Wert

$$\lim_{\delta \rightarrow \infty} \mu(x_2^*) = b_{1,2}$$

zustrebt. Als Asymptote kann somit der Wert $\min\{b_1, b_2\}$ betrachtet werden. Abbildung 3.3 zeigt die Ergebnisse der Untersuchung des Distanzmaßes μ für $b_1 \neq b_2$. Hierbei wurde der Wert b_1 konstant gleich Eins gewählt, und der Wert b_2 variiert.

Aus dieser Abbildung ist ersichtlich, dass mit größer werdenden Werten für b_2 die asymptotische Darstellung schlechter wird. Es können jedoch folgende Schlüsse gezogen werden. Für hinreichend große Werte δ strebt das Distanzmaß gegen den Wert $\min\{b_1, b_2\}$. Für kleine Werte der Elemente von \mathbf{b} und kleine δ wird das Maß fast ausschließlich durch den Abstand der Eigenwerte δ bestimmt.

Zusammenfassend kann aus der erarbeiteten asymptotischen Darstellung geschlossen werden, dass, wenn die Eigenwerte nahe beieinander liegen, das Distanzmaß klein und somit die Steuerbarkeit schlecht ist. Liegen die Eigenwerte hinreichend weit von einander entfernt, sind die Werte im Eingangsvektor \mathbf{b} für die Güte der Steuerbarkeit verantwortlich. Ist das kleinste Ele-

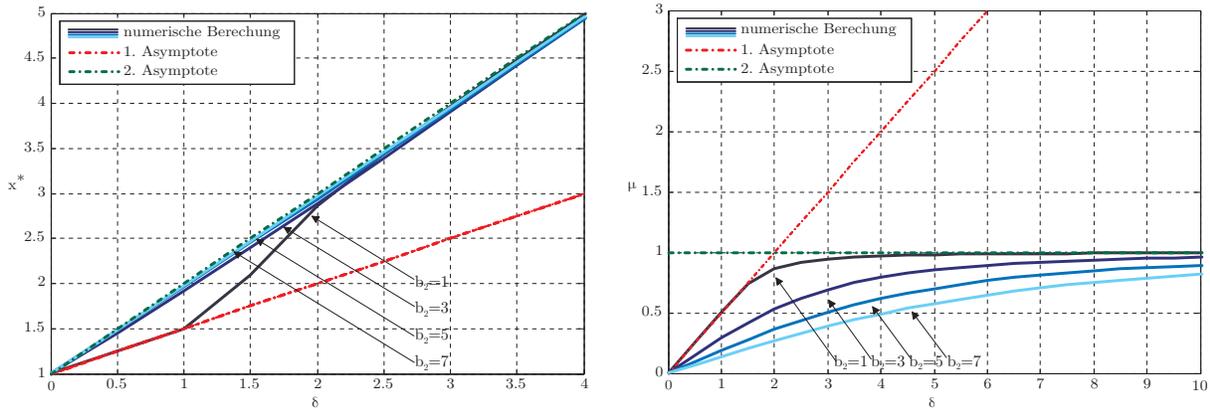


Abbildung 3.3: Numerische und asymptotische Lösung für das Minimum x^* (links) und das Distanzmaß μ (rechts)

ment in \mathbf{b} klein, so führt dies zu einer kleinen Maßzahl und deutet somit auf schlechte Steuerbarkeit hin.

3.1.2.2 Konjugiert komplexe Eigenwerte

Es sei⁴

$$\lambda_1 = \bar{\lambda}_2 =: \lambda = \sigma + j\omega \quad \text{und} \quad b_1 = \bar{b}_2 =: b.$$

Damit erhält man mit $\mathbf{M} = [s\mathbf{E} - \mathbf{A}, \mathbf{b}]$ die Matrix

$$\mathbf{M}\mathbf{M}^H = \begin{bmatrix} \alpha_1 & b^2 \\ \bar{b}^2 & \alpha_2 \end{bmatrix}$$

mit

$$\alpha_1 := (x - \sigma)^2 + (y - \omega)^2 + |b|^2 \quad \text{und} \quad \alpha_2 := (x - \sigma)^2 + (y + \omega)^2 + |b|^2.$$

Für die Eigenwerte $\Lambda_{1,2}$ der Matrix $\mathbf{M}\mathbf{M}^H$ gilt

$$\Lambda_{1,2} = (x - \sigma)^2 + y^2 + \omega^2 + |b|^2 \pm \sqrt{|b|^4 + 4y^2\omega^2}.$$

Durch Nullsetzen der Ableitung nach x ergibt sich

$$\frac{\partial \Lambda_{1,2}}{\partial x} = 2x - 2\sigma \stackrel{!}{=} 0 \quad \curvearrowright \quad x^* = \sigma. \quad (3.12)$$

Analog erhält man durch die Ableitung nach y die drei Lösungen $y_1^* = 0$ und

$$y_{2,3}^* = \pm \frac{1}{2\omega} \sqrt{4\omega^4 - |b|^4}. \quad (3.13)$$

Dadurch ergeben sich folgende mögliche Lösungen für das Distanzmaß μ_C : Ist an der Stelle $s_1^* = x^* + jy_1^*$ das globale Minimum von Gleichung (3.3), gilt für das Distanzmaß

$$\mu_1 := \mu_C|_{s=s_1^*} = \omega.$$

⁴Mit \bar{z} wird die zu z konjugiert komplexe Zahl symbolisiert.

Da $\sigma_{\min}([s\mathbf{E} - \mathbf{A}, \mathbf{B}]) = \sigma_{\min}([\bar{s}\mathbf{E} - \mathbf{A}, \mathbf{B}])$ gilt, liefern die Werte $s_{2,3}^* = x^* + jy_{2,3}^*$ für das Distanzmaß μ_C das identische Ergebnis

$$\mu_2 := \mu_C|_{s=s_{2,3}^*} = \frac{1}{2} \left| \frac{b}{\omega} \right| \sqrt{4\omega^2 - |b|^2}. \quad (3.14)$$

Die Lösung μ_2 ist nur dann gültig, wenn sie reell ist. D.h., es muss die Ungleichung

$$4\omega^2 \geq |b|^2$$

erfüllt sein. Ist $|b|^2 > 4\omega^2$, womit nur s_1^* gültig sein kann, dann gilt für das Distanzmaß μ_C die Lösung μ_1 . Anderenfalls sind μ_1 und μ_2 mögliche Lösungen für μ_C . Da jedoch für diesen Fall immer $\mu_2 < \mu_1$, gilt für $|b|^2 \leq 4\omega^2$ die Lösung μ_2 .

Auffällig ist, dass beide Lösungen μ_1 und μ_2 unabhängig von Realteil des Eigenwertes der Matrix \mathbf{A} sind. Für hinreichend große Werte für b wird die Lage des Minimums allein durch den Realteil des Eigenwertes von \mathbf{A} bestimmt, während das Distanzmaß μ_C selbst nur vom Imaginärteil des Eigenwertes abhängig ist.

3.1.2.2.1 Distanzmaß μ_R für Systeme 2. Ordnung mit komplexen Eigenwerten Werden bei der Bestimmung des Distanzmaßes nur reelle Störungen ($\delta\mathbf{A} \in \mathbb{R}^{n \times n}$ und $\delta\mathbf{B} \in \mathbb{R}^{n \times m}$) berücksichtigt, so entspricht das Maß dem Minimum von Gleichung (3.3) auf der reellen Achse. Von den drei zuvor ermittelten lokalen Minima liegt nur eines auf der reellen Achse, und somit gilt für das Minimum $s^* = \sigma$ und $\mu_R = \mu_1 = \omega$. Das Distanzmaß μ_R für Systeme 2. Ordnung in Modalform ist folglich nur vom Imaginärteil der Eigenwerte der Matrix \mathbf{A} abhängig. Der Eingangsvektor \mathbf{b} hat keinen Einfluss auf das Distanzmaß μ_R .

Zusammenfassend können folgende Ergebnisse festgehalten werden:

Zusammenfassung (Systeme 2. Ordnung in Modalform):

- $\lambda(\mathbf{A}) \in \mathbb{R}$:
 - $\mu_C = \mu_R$
- $\lambda(\mathbf{A}) \in \mathbb{C}$ und komplexe Störungen (Distanzmaß μ_C):
 - $|b|^2 > 4\omega^2$
 - Lage des Minimums $s^* = \sigma \quad \Rightarrow \mu_C = \mu_R$
 - Steuerbarkeitsmaß $\mu_C = \omega$
 - $|b|^2 \leq 4\omega^2$
 - Lage des Minimums $s^* = \sigma \pm \frac{1}{2\omega} \sqrt{4\omega^4 - |b|^4}$
 - Steuerbarkeitsmaß $\mu_C = \frac{1}{2} \left| \frac{b}{\omega} \right| \sqrt{4\omega^2 - |b|^2}$
- $\lambda(\mathbf{A}) \in \mathbb{C}$ und reelle Störungen (Distanzmaß μ_R):
 - Lage des Minimums $s^* = \sigma$

- Steuerbarkeitsmaß $\mu_R = \omega$
- \Rightarrow der Eingangsvektor \mathbf{b} hat keinen Einfluss auf das Maß

3.1.3 Systeme höherer Ordnung

Für Systeme höherer Ordnung stellt sich die Ermittlung einer analytischen Lösung erheblich schwieriger dar. Durch den höheren Grad der durch das Ableiten von σ nach x bzw. y entstehenden Polynome erschwert sich das Auffinden von Nullstellen. Ebenfalls werden die Ungleichungen, die bei mehreren Nullstellen zu betrachten sind, wesentlich komplizierter.

Trotzdem kann eine Reihe nützlicher Aussagen getroffen werden: Besitzt die Systemmatrix \mathbf{A} bestimmte Eigenschaften, dann kann gezeigt werden, dass für das Distanzmaß $\mu_C = \mu_R$ gilt. Dadurch vereinfacht sich die Berechnung von μ bzw. die Suche nach dem Minimum erheblich, da nur noch entlang der reellen Achse gesucht wird. Damit können bei der numerischen Bestimmung des Maßes die in Abschnitt 3.2.2 vorgestellten Algorithmen zur Ermittlung von μ_R eingesetzt werden. Diese sind i.A. einfacher und vor allem verringert sich der Aufwand der numerischen Berechnung (siehe 3.2.6).

Ist die Systemmatrix \mathbf{A} symmetrisch, dann gilt $\mu_C = \mu_R$. Dies folgt direkt aus dem in Ungleichung (3.6) definierten Rechteck, in welchem die Minima liegen. In diesem Fall entartet dieses Rechteck zu einer Geraden auf der reellen Achse.

Da Singulärwerte invariant gegenüber einer unitären Zustandstransformation sind, folgt daraus, dass auch das Distanzmaß gegenüber einer solchen Transformation invariant ist. Im Allgemeinen ist die Transformation auf Modalform jedoch nicht unitär. Dann gilt (vgl. [20]): Für

$$\mathbf{A} = \mathbf{X}_R \hat{\mathbf{A}} \mathbf{X}_R^{-1} \quad \text{mit} \quad \|\mathbf{X}_R\| = 1$$

erhält man wegen

$$\mathbf{X}_R^{-1} [s\mathbf{E} - \mathbf{A}, \mathbf{B}] \begin{bmatrix} \mathbf{X}_R & \mathbf{0} \\ \mathbf{0} & \mathbf{E} \end{bmatrix} = [s\mathbf{E} - \hat{\mathbf{A}}, \hat{\mathbf{B}}]$$

den Zusammenhang

$$\mu(\mathbf{A}, \mathbf{B}) \geq \frac{1}{K(\mathbf{X}_R)} \mu(\hat{\mathbf{A}}, \hat{\mathbf{B}}).$$

Hierbei ist $K(\mathbf{X}_R) := \|\mathbf{X}_R\| \|\mathbf{X}_R^{-1}\|$ die sogenannte Konditionszahl der Matrix der Rechtseigenvektoren der Systemmatrix \mathbf{A} . Für die Konditionszahl gilt

$$K(\mathbf{X}_R) = \frac{\sigma_{\max}(\mathbf{X}_R)}{\sigma_{\min}(\mathbf{X}_R)}.$$

Durch diese Beziehung kann eine untere Schranke für das Distanzmaß angegeben werden, indem das Maß des Systems in Modalform mit dem Verhältnis des größten zum kleinsten Singulärwerts der Matrix \mathbf{X}_R skaliert wird:

$$\mu(\mathbf{A}, \mathbf{B}) \geq \frac{\sigma_{\min}(\mathbf{X}_R)}{\sigma_{\max}(\mathbf{X}_R)} \mu(\hat{\mathbf{A}}, \hat{\mathbf{B}}).$$

Im Folgenden wird eine obere Schranke für das Distanzmaß μ_C in Abhängigkeit der Eigenwerte der Matrix \mathbf{A} und des Eingangsvektors \mathbf{b} angegeben. Dadurch erhält man auch für Systeme höherer Ordnung ein besseres Verständnis für das Distanzmaß und somit für das Steuerbarkeitsverhalten des Systems. Für das Distanzmaß gilt der Zusammenhang

$$\mu_C \leq \min_i \{\sigma(\lambda_i)\} \quad i = 1, \dots, n.$$

Hierbei sind λ_i die Eigenwerte von \mathbf{A} und $\sigma(\lambda_i)$ der kleinste Singulärwert der Matrix

$$\mathbf{M}(\lambda_i) := [\lambda_i \mathbf{E} - \mathbf{A}, \mathbf{b}].$$

In Abschnitt 3.2.3 wird gezeigt, dass diese Schranke eine gute Abschätzung für das Distanzmaß liefert.

Der kleinste Singulärwert kann nun folgendermaßen abgeschätzt werden:

$$\sigma(\lambda_i) = \frac{1}{\|\mathbf{M}^\dagger(\lambda_i)\|_2} \leq \|\mathbf{M}(\lambda_i)\|_2.$$

Hierbei wird mit \mathbf{M}^\dagger die Pseudoinverse der Matrix \mathbf{M} bezeichnet.

Die Spektralnorm $\|\cdot\|_2$ kann wiederum durch die Wurzel aus dem Produkt der Spaltensummennorm $\|\cdot\|_1$ und der Zeilensummennorm $\|\cdot\|_\infty$ abgeschätzt werden

$$\|\mathbf{M}(\lambda_i)\|_2 \leq \sqrt{\|\mathbf{M}(\lambda_i)\|_1 \|\mathbf{M}(\lambda_i)\|_\infty}.$$

Für ein in Modalform vorliegendes System

$$(\mathbf{A} = \text{diag}(\lambda_i), \mathbf{b})$$

hat die Matrix $\mathbf{M}(\lambda_i)$ die Struktur⁵

$$\mathbf{M}(\lambda_i) = \begin{bmatrix} \lambda_i - \lambda_1 & 0 & \cdots & 0 & b_1 \\ 0 & \lambda_i - \lambda_2 & & \vdots & \vdots \\ \vdots & & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & \lambda_i - \lambda_n & b_n \end{bmatrix}.$$

Die Normen können nun wie folgt berechnet werden: Wir definieren den Abstand zwischen zwei Eigenwerten der Matrix $\hat{\mathbf{A}}$

$$\delta_{ij} = |\lambda_i - \lambda_j| \quad i, j = 1, \dots, n$$

und den maximalen Abstand des Eigenwerts λ_i zu den übrigen Eigenwerten

$$\delta_{i,\max} = \max_{j, j \neq i} \{\delta_{ij}\}.$$

Mit diesen Abkürzungen gilt für die Spaltensummennorm der Matrix $\mathbf{M}(\lambda_i)$

$$\|\mathbf{M}(\lambda_i)\|_1 = \max\{\delta_{i,\max}, \sum_{i=1}^n |b_i|\}.$$

⁵Die durchgeführten Überlegungen lassen sich sehr einfach auf Mehrgrößensysteme übertragen. Aus Gründen der Übersichtlichkeit ist hier nur der Eingrößenfall angeführt.

Die Zeilensummennorm dieser Matrix lässt sich berechnen durch

$$\|\mathbf{M}(\lambda_i)\|_\infty = \max_{j, j \neq i} \{\delta_{ij}, |b_j|\}.$$

Diese obere Schranke für das Distanzmaß μ verdeutlicht, dass sowohl die Elemente im Vektor \mathbf{b} als auch der Abstand zwischen den Eigenwerten der Systemmatrix den Wert μ beeinflussen. Da für die Schranke das Minimum

$$\mu_C \leq \min_i \|\mathbf{M}(\lambda_i)\|_2$$

gebildet wird, gilt für hinreichend kleine Elemente im Vektor \mathbf{b} , dass der kleinste Maximalabstand zwischen den Eigenwerten entscheidend ist. Sind die Abstände zwischen den Eigenwerten klein, so wird die obere Schranke von den Elementen des Vektors \mathbf{b} bestimmt. Sind diese klein, dann ist die obere Schranke und somit auch das Distanzmaß μ klein. Somit kommt man, wie schon bei der Betrachtung der Systeme 1. und 2. Ordnung, zu dem Ergebnis, dass das Distanzmaß μ klein bzw. die Steuerbarkeit schlecht ist, wenn die Eigenwerte der Systemmatrix \mathbf{A} nahe beieinanderliegen oder die Elemente des Eingangsvektors \mathbf{b} klein sind.

Für Systeme allgemeiner Form liefert die Betrachtung dieser Schranke kein so einfaches Ergebnis.

3.2 Numerische Bestimmung

Um für Systeme höherer Ordnung das Distanzmaß zu ermitteln, muss auf eine numerische Bestimmung zurückgegriffen werden. Prinzipiell können die hier vorgestellten Verfahren zur Bestimmung von μ_C bzw. von μ_R in folgende Gruppen unterteilt werden:

Ansatz: Optimierungsproblem

Einige der Verfahren gehen direkt von dem Optimierungsproblem (z.B. nach Gleichung (3.3)) aus, um durch geeignete Suchverfahren das Minimum näherungsweise zu ermitteln. D.h., es wird in dem über der komplexen Ebene aufgespannten 3D-Gebirge nach Minima gesucht. Hierbei gibt es zwei unterschiedliche Vorgehensweisen. Eine ist die direkte Suche nach dem globalen Minimum. Die zweite Möglichkeit ist die Suche nach lokalen Minima mit einer zusätzlichen Überprüfung, ob es sich bei dem gefundenen Minimum um das Globale handelt. Bei der Bestimmung von μ_R wird sich zeigen, dass durch die Beschränkung der Suche auf die reelle Achse die Algorithmen im Wesentlichen einer Vereinfachung derer zur Bestimmung von μ_C entsprechen.

Ansatz: Störungstheorie

Eine andere Vorgangsweise ergibt sich, wenn das Problem aus Sicht der Störungstheorie betrachtet wird. Hierbei werden die Matrizen $\delta\mathbf{A}$ und $\delta\mathbf{B}$ als Störung interpretiert und mit Hilfe der aus der Störungstheorie bekannten Theoreme die Eigenwerte bzw. Singulärwerte des gestörten Systems untersucht.

Abbildung 3.4 gibt einen Überblick über die in dieser Arbeit vorgestellten Verfahren zur Abschätzung von μ_C bzw. μ_R . Der Algorithmus nach WICKS und DeCARLO liefert einen Schätzwert, wohingegen alle übrigen Ansätze Schranken für das Distanzmaß μ liefern. Die mit (*) gekennzeichneten Ansätze liefern nur eine obere Schranke, während die übrigen sowohl eine unter als auch eine obere Schranke für das Distanzmaß bestimmen.

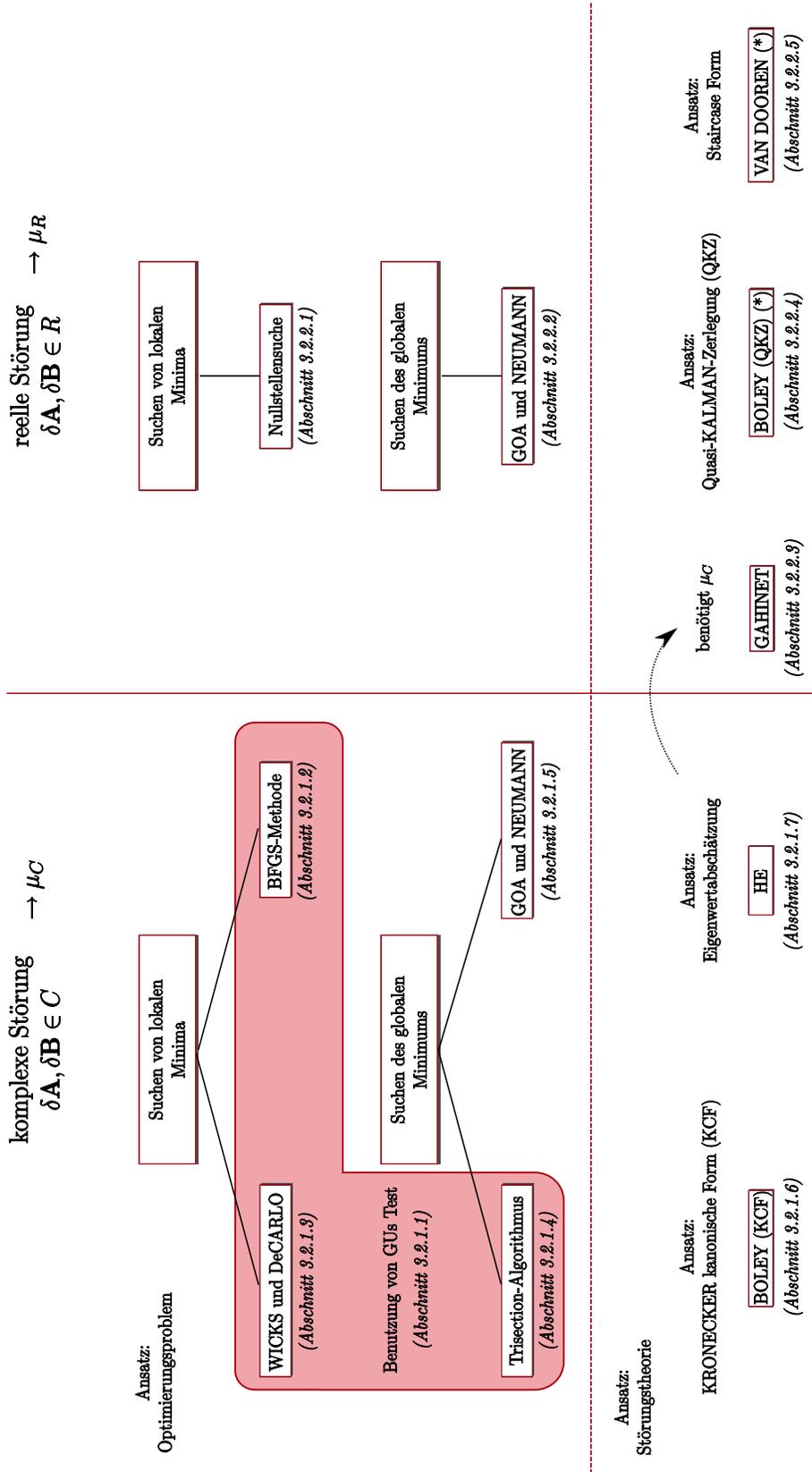


Abbildung 3.4: Übersicht über die numerische Berechnung von Schranken für das Distanzmaß μ

3.2.1 Numerische Bestimmung der Größe μ_C

Im Folgendem werden vier Methoden vorgestellt, die basierend auf dem Optimierungsproblem Schranken für das Distanzmaß μ_C ermitteln (Abschnitt 3.2.1.2 bis 3.2.1.5). Beim Trisection-Algorithmus, sowie bei der Methode nach GOA und NEUMANN wird direkt nach dem globalen Minimum gesucht. Bei der BFGS-Methode und dem Algorithmus nach WICKS und DECARLO werden nach lokalen Minima gesucht. Die notwendige Überprüfung, ob es sich um das globale Minimum handelt, erfolgt mit Hilfe von GUs Test (Abschnitt 3.2.1.1). In den Abschnitten 3.2.1.6 und 3.2.1.7 werden zwei Methoden vorgestellt, bei denen die Störungstheorie zur Bestimmung von Schranken für μ_C genutzt wird.

3.2.1.1 GUs Test

Werden zwei positive Zahlen δ_1 und δ_2 vorgegeben, liefert der Test eine Aussage über den Wert von μ_C in Relation zu δ_1 bzw. δ_2 :

$$\mu_C \leq \delta_1 \quad \text{oder} \quad \mu_C > \delta_2.$$

Hierfür wird von folgendem Theorem ausgegangen [18]:

Theorem 5: Es sei δ_1 eine Zahl mit der Eigenschaft

$$\delta_1 > \mu_C(\mathbf{A}, \mathbf{B}).$$

Für alle Werte von η mit

$$0 \leq \eta \leq 2(\delta_1 - \mu_C(\mathbf{A}, \mathbf{B}))$$

existieren *mindestens* zwei Paare reeller Zahlen α und β , so dass

$$\sigma_{\min}([\mathbf{A} - (\alpha + \beta i)\mathbf{E}, \mathbf{B}]) = \delta_1 \quad \text{und} \quad (3.15a)$$

$$\sigma_{\min}([\mathbf{A} - (\alpha + \eta + \beta i)\mathbf{E}, \mathbf{B}]) = \delta_1 \quad (3.15b)$$

gilt. Existiert *kein* Paar (α, β) , so gilt

$$\eta > 2(\delta_1 - \mu_C(\mathbf{A}, \mathbf{B})),$$

und daraus folgt

$$\mu_C(\mathbf{A}, \mathbf{B}) > \delta_1 - \frac{\eta}{2} =: \delta_2.$$

Zum leichteren Verständnis dieses Theorems dient folgende graphische Darstellung:

Es sei $\delta_1 \geq \mu_C(\mathbf{A}, \mathbf{B})$. Für

$$\sigma_{\min}([\mathbf{A} - (\alpha + \beta i)\mathbf{E}, \mathbf{B}]) = \delta_1$$

muss gelten

$$\det([\mathbf{A} - (\alpha + \beta i)\mathbf{E}][\mathbf{A}^H - (\alpha - \beta i)\mathbf{E}] + \mathbf{B}\mathbf{B}^H - \delta_1^2\mathbf{E}) = 0. \quad (3.16)$$

Die Lösung dieser Gleichung entspricht einer endlichen Anzahl geschlossener Kurven in der $\alpha\beta$ -Ebene. Diese Kurven entsprechen den Höhengichtlinien mit dem Wert δ_1 des durch Gleichung (3.3) aufgespannten Gebirges. Abbildung 3.5 zeigt eine mögliche Lösung in der $\alpha\beta$ -Ebene. Alle Punkte (α, β) auf dem Rand $\partial\mathcal{G}$ des Gebietes \mathcal{G} erfüllen Gleichung (3.16). Die

Lösung (α^*, β^*) , die Relation (3.3) zu einem Minimum macht, muss aufgrund der Annahme $\delta_1 \geq \mu_C(\mathbf{A}, \mathbf{B})$ im Inneren einer dieser Kurven liegen, d.h. $(\alpha^*, \beta^*) \in \mathcal{G}$. Die beiden Punkte $P_1 = (\alpha^* - \eta_1, \beta^*)$ und $P_2 = (\alpha^* + \eta_2, \beta^*)$ werden nun so gewählt, dass $P_1, P_2 \in \partial\mathcal{G}$ gilt. Damit gilt

$$\sigma(\alpha^* - \eta_1 + \beta^*i) = \sigma(\alpha^* + \eta_2 + \beta^*i) = \delta_1.$$

Unter Benutzung des Theorems nach MIRSKY (siehe Anhang A.4.2) gilt folgende Ungleichung

$$\delta_1 - \mu_C(\mathbf{A}, \mathbf{B}) \leq \|[\mathbf{A} - (\alpha^* - \eta_1 + \beta^*i)\mathbf{E}, \mathbf{B}] - [\mathbf{A} - (\alpha^* + \beta^*i)\mathbf{E}, \mathbf{B}]\|.$$

Daraus folgt

$$\eta_1 \geq \delta_1 - \mu_C(\mathbf{A}, \mathbf{B}) \quad \text{und analog} \quad \eta_2 \geq \delta_1 - \mu_C(\mathbf{A}, \mathbf{B}). \quad (3.17)$$

Das Gebiet \mathcal{G} wird nun um den Wert η nach links verschoben. Das so entstandene Gebiet $\hat{\mathcal{G}}$ ist in Abbildung 3.6 dargestellt. Liegt die Verschiebung η im Intervall $[0, 2(\delta_1 - \mu_C(\mathbf{A}, \mathbf{B}))]$, dann ist wegen (3.17) sichergestellt, dass sich die beiden Kurven $\partial\mathcal{G}$ und $\partial\hat{\mathcal{G}}$ schneiden. Für die Schnittpunkte $\tilde{P}_i = (\tilde{\alpha}_i, \tilde{\beta}_i)$ gilt $\tilde{P}_i \in \partial\mathcal{G}$ und $(\tilde{\alpha}_i + \eta, \tilde{\beta}_i) \in \partial\hat{\mathcal{G}}$. Damit existieren für $\eta \in [0, 2(\delta_1 - \mu_C(\mathbf{A}, \mathbf{B}))]$ mindestens zwei Punkte \tilde{P}_i , welche die Gleichungen (3.15) aus Theorem 5 erfüllen.

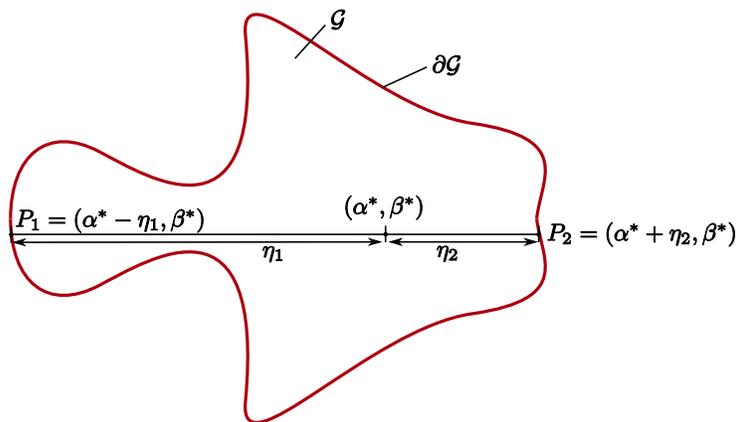


Abbildung 3.5: Mögliche Lösung von Gleichung (3.16)

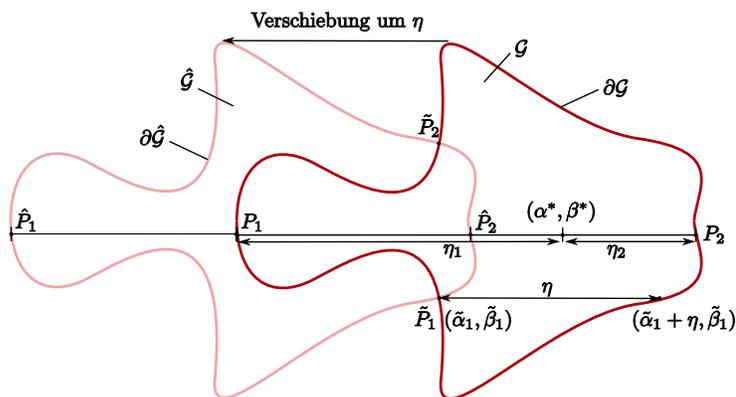


Abbildung 3.6: Um η verschobene Lösung

Die Existenz eines in Theorem 5 beschriebenen Paares (α, β) kann nach [18] folgendermaßen festgestellt werden: Da δ_1 ein Singulärwert von $[\mathbf{A} - (\alpha + \beta i)\mathbf{E}, \mathbf{B}]$ ist, existieren Vektoren $\mathbf{g}, \mathbf{h}, \mathbf{z}$ mit der Eigenschaft:

$$(\mathbf{A} - (\alpha + \beta i)\mathbf{E}, \mathbf{B}) \begin{pmatrix} \mathbf{g} \\ \mathbf{h} \end{pmatrix} = \delta_1 \mathbf{z} \quad \text{und} \quad \begin{pmatrix} \mathbf{A}^H - (\alpha - \beta i)\mathbf{E} \\ \mathbf{B}^H \end{pmatrix} \mathbf{z} = \delta_1 \begin{pmatrix} \mathbf{g} \\ \mathbf{h} \end{pmatrix}. \quad (3.18)$$

Obige Gleichungen können folgendermaßen umgeformt werden:

$$\underbrace{\begin{pmatrix} \mathbf{A} - \alpha\mathbf{E} & \mathbf{B}\mathbf{Q}_{22} - \delta_1\mathbf{Q}_{12} \\ -\delta_1\mathbf{Q}_{12}^{-1} & \mathbf{Q}_{12}^{-1}(\mathbf{A}^H - \alpha\mathbf{E})\mathbf{Q}_{12} \end{pmatrix}}_{=: \mathbf{H}(\alpha, \delta_1)} \begin{pmatrix} \mathbf{g} \\ \mathbf{h}_1 \end{pmatrix} = \beta i \begin{pmatrix} \mathbf{g} \\ \mathbf{h}_1 \end{pmatrix}. \quad (3.19)$$

Die Matrizen \mathbf{Q}_{11} bis \mathbf{Q}_{22} ergeben sich aus der QR-Zerlegung

$$\begin{pmatrix} \mathbf{B} \\ -\delta_1\mathbf{E} \end{pmatrix} = \begin{pmatrix} \mathbf{Q}_{11} & \mathbf{Q}_{12} \\ \mathbf{Q}_{21} & \mathbf{Q}_{22} \end{pmatrix} \begin{pmatrix} \mathbf{R} \\ \mathbf{0} \end{pmatrix}.$$

Der Vektor \mathbf{h}_1 ergibt sich aus

$$\mathbf{h}_1 = \mathbf{Q}_{22}^H \mathbf{h}.$$

Analoges gilt auch für Gleichung (3.15b), und man erhält

$$\underbrace{\begin{pmatrix} \mathbf{A} - (\alpha + \eta)\mathbf{E} & \mathbf{B}\mathbf{Q}_{22} - \delta\mathbf{Q}_{12} \\ -\delta_1\mathbf{Q}_{12}^{-1} & \mathbf{Q}_{12}^{-1}(\mathbf{A}^H - (\alpha + \eta)\mathbf{E})\mathbf{Q}_{12} \end{pmatrix}}_{\mathbf{H}(\alpha + \eta, \delta_1)} \begin{pmatrix} \mathbf{g} \\ \mathbf{h}_1 \end{pmatrix} = \beta i \begin{pmatrix} \mathbf{g} \\ \mathbf{h}_1 \end{pmatrix}. \quad (3.20)$$

Damit die Gleichungen (3.15a) und (3.15b) mindestens ein reelles Lösungspaar (α, β) besitzen, müssen die Gleichungen (3.19) und (3.20) einen gemeinsamen rein imaginären Eigenwert βi haben. Um dies zu erfüllen, muss die SYLVESTER-Gleichung

$$\mathbf{H}(\alpha, \delta_1)\mathbf{X} - \mathbf{X}\mathbf{H}^H(\alpha + \eta, \delta_1) = 0$$

eine nicht triviale Lösung besitzen. Diese Gleichung führt nach einigen mühevollen Umformungen (siehe [18]) zu einem generalisierten Eigenwertproblem der Form

$$\mathcal{A}\mathbf{v} = 2\alpha\mathcal{B}\mathbf{v}.$$

Mit Hilfe dieser Erkenntnisse kann nun GUs Test gemäß folgendem Algorithmus durchgeführt werden:

Algorithmus 3.1: GUs Test

1. Berechne die reellen Eigenwerte von $\mathcal{A}\mathbf{v} = 2\alpha\mathcal{B}\mathbf{v}$.
2. Überprüfe für jeden reellen Eigenwert α , ob die Matrizen $\mathbf{H}(\alpha, \delta_1)$ und $\mathbf{H}(\alpha + \eta, \delta_1)$ einen gemeinsamen rein imaginären Eigenwert βi besitzen.

Wird entweder kein reelles α oder kein reelles β gefunden, folgt daraus

$$\mu(\mathbf{A}, \mathbf{B}) > \delta_2.$$

Ansonsten gilt

$$\mu(\mathbf{A}, \mathbf{B}) \leq \delta_1.$$

3.2.1.2 BFGS-Methode unter Verwendung von GUs Test

Bei dem in [6] vorgestellten Algorithmus wird mit Hilfe der Methode nach Broyden-Fletcher-Goldfarb-Shanno (BFGS) ein lokales Minimum gesucht und mit GUs Test bestimmt, ob es sich um das globale Minimum handelt. Die BFGS-Methode ist ein sogenanntes Quasi-NEWTON-Verfahren. Beim NEWTON-Verfahren werden die Nullstellen einer Funktion f gesucht, wobei f durch eine quadratische Funktion approximiert wird.

Durch eine TAYLOR-Reihenentwicklung von $f(\mathbf{x} + \Delta\mathbf{x})$ um \mathbf{x} ergibt sich

$$f(\mathbf{x} + \Delta\mathbf{x}) = f(\mathbf{x}) + \Delta\mathbf{x}^T \nabla f(\mathbf{x}) + \frac{1}{2} \Delta\mathbf{x}^T \mathbf{H}(f)(\mathbf{x}) \Delta\mathbf{x} + \dots$$

mit dem Gradienten von f

$$\nabla f(\mathbf{x}) = \left(\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} \right)^T = \begin{pmatrix} \frac{\partial f(\mathbf{x})}{\partial x_1} \\ \frac{\partial f(\mathbf{x})}{\partial x_2} \\ \vdots \\ \frac{\partial f(\mathbf{x})}{\partial x_n} \end{pmatrix}$$

und mit $\mathbf{H}(f)$ der sogenannten HESSE-Matrix

$$\mathbf{H}(f) := \begin{pmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_n} \end{pmatrix}.$$

Wir betrachten nur eine „kleine“ Umgebung von \mathbf{x} , sodass Terme höherer Ordnung vernachlässigt werden können. Die Funktion f verhält sich in dieser Umgebung annähernd quadratisch. Im k -ten Schritt des Verfahrens wird die Funktion f in der Umgebung von \mathbf{x}_k durch eine quadratische Funktion approximiert. Für die Berechnung des nächsten Schrittes \mathbf{x}_{k+1} wird das Extremum der quadratischen Funktion wie folgt ermittelt

$$\Delta\mathbf{x}_k = -(\mathbf{H}(f)(\mathbf{x}_k))^{-1} \nabla f(\mathbf{x}_k).$$

Damit ergibt sich für den $(k + 1)$ -ten Schritt

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \Delta\mathbf{x}_k.$$

Es existieren eine Vielzahl von Verfahren, die, um den Aufwand zu reduzieren, die Inverse HESSE-Matrix approximieren. Diese werden in der Literatur oft als Quasi-NEWTON-Verfahren bezeichnet (siehe z.B. [47]).

Unter der Voraussetzung, dass der kleinste Singulärwert in Relation (3.3) einfach auftritt, kann der Gradient der zu minimierenden Funktion $\sigma(s)$ mit Hilfe der in Gleichung (3.4) eingeführten Größe $f(s)$ analytisch angegeben werden. ELSNER und HE haben gezeigt, dass für die Ableitung von $\sigma(s)$ nach x und y gilt

$$\frac{\partial \sigma(s)}{\partial x} = -\operatorname{Re}\{f\} \quad \text{bzw.} \quad \frac{\partial \sigma(s)}{\partial y} = -\operatorname{Im}\{f\}.$$

Auch die HESSE-Matrix kann für diesen Fall durch Bestimmung der 2. Ableitung analytisch ermittelt werden (siehe [11]).

Für die hier durchgeführten Berechnungen wird ein Quasi-NEWTON-Verfahren, die so genannte BFGS-Methode, verwendet, um das Invertieren dieser Matrix zu umgehen. Diese Methode benötigt nur den Gradienten der zu minimierenden Funktion und konvergiert allerdings ein wenig langsamer als das NEWTON-Verfahren.

Wenn ein Minimum ermittelt wurde, kann mit GUs Test festgestellt werden, ob es sich um das globale Minimum handelt: Das ermittelte Minimum sei mit $\tilde{\mu}_C$ bezeichnet. Man wählt

$$\delta_1 = \tilde{\mu}_C \left(1 - \frac{1}{2}\epsilon\right) \quad \text{und} \quad \delta_2 = \tilde{\mu}_C (1 - \epsilon)$$

mit einem kleinen, positiven Wert ϵ . Liefert GUs Test das Ergebnis $\mu_C \leq \delta_1$, handelt es sich um ein lokales Minimum. Erhält man als Ergebnis $\mu_C > \delta_2$, dann handelt es sich bei dem gefundenen Minimum $\tilde{\mu}_C$ um das globale Minimum (siehe Abbildung 3.7).

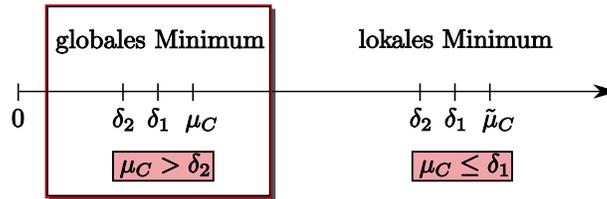


Abbildung 3.7: Detektion des globalen Minimums mit Hilfe von GUs Test

3.2.1.3 Algorithmus nach WICKS und DeCARLO

EISING hat in [10] gezeigt, dass das Optimierungsproblem

$$\begin{aligned} \mu^2 &= \inf_{\delta\mathbf{A}, \delta\mathbf{B}} \{d_A \|\delta\mathbf{A}\|^2 + d_B \|\delta\mathbf{B}\|^2\} \\ &= \inf_{\mathbf{q} \in \mathbb{C}^n, \|\mathbf{q}\|=1} \{d_A \mathbf{q}^H \mathbf{A} (\mathbf{E} - \mathbf{q}\mathbf{q}^H) \mathbf{A}^T \mathbf{q} + d_B \mathbf{q}^H \mathbf{B} \mathbf{B}^T \mathbf{q}\} \end{aligned} \quad (3.21)$$

für $d_A = d_B = 1$ dem Optimierungsproblem (3.3) äquivalent ist. Die reellen und positiven Parameter d_A und d_B ermöglichen eine unterschiedliche Gewichtung der Störungen auf die Matrizen \mathbf{A} und \mathbf{B} . Für die Herleitung obiger Beziehung geht EISING von folgendem Theorem aus:

Theorem 6: Das System $(\mathbf{A} + \delta\mathbf{A}, \mathbf{B} + \delta\mathbf{B})$ ist genau dann nicht steuerbar, wenn ein von Null verschiedener Vektor \mathbf{q}^H existiert, so dass gilt

$$\mathbf{q}^H (\mathbf{A} + \delta\mathbf{A}) = s \mathbf{q}^H \quad \mathbf{q}^H (\mathbf{B} + \delta\mathbf{B}) = \mathbf{0}.$$

EISING nimmt nun an, dass die Matrizen \mathbf{A} und \mathbf{B} sowie der Vektor \mathbf{q} gegeben seien und sucht nach den beiden Matrizen $\delta\mathbf{A}$ und $\delta\mathbf{B}$, so dass obiges Theorem für irgendein $s \in \mathbb{C}$ erfüllt ist. Dadurch erhält man für das $\delta\mathbf{A}$ und $\delta\mathbf{B}$ mit der kleinsten Norm

$$\delta\mathbf{A} = \frac{\mathbf{q} \left(-\mathbf{q}^H \mathbf{A} + \frac{\mathbf{q}^H \mathbf{A} \mathbf{q}}{\mathbf{q}^H \mathbf{q}} \mathbf{q}^H \right)}{\mathbf{q}^H \mathbf{q}} \quad \text{und} \quad \delta\mathbf{B} = -\frac{\mathbf{q} \mathbf{q}^H \mathbf{B}}{\mathbf{q}^H \mathbf{q}}.$$

Durch die Berechnung von $\|\delta\mathbf{A}\|^2$ und $\|\delta\mathbf{B}\|^2$ mit $\mathbf{q}^H\mathbf{q} = 1$ erhält man das Optimierungsproblem (3.21).

WICKS und DeCARLO kamen mit dem Ansatz der KALMAN-Zerlegung (siehe Anhang A.2) auf das gleiche Ergebnis. Wie aus der KALMAN-Zerlegung bekannt, gilt bei einer Partitionierung der Matrizen

$$\tilde{\mathbf{A}} = \mathbf{P}\mathbf{A}\mathbf{P}^T = \begin{bmatrix} \tilde{\mathbf{A}}_{11} & \tilde{\mathbf{A}}_{12} \\ \tilde{\mathbf{A}}_{21} & \tilde{\mathbf{A}}_{22} \end{bmatrix} \quad \text{und} \quad \tilde{\mathbf{B}} = \mathbf{P}\mathbf{B} = \begin{bmatrix} \tilde{\mathbf{B}}_1 \\ \tilde{\mathbf{B}}_2 \end{bmatrix}, \quad (3.22)$$

dass das System (\mathbf{A}, \mathbf{B}) genau dann *nicht* steuerbar ist, wenn $\tilde{\mathbf{A}}_{21}$ und $\tilde{\mathbf{B}}_2$ Nullmatrizen sind. Für die kleinst mögliche Störung, die das System nicht steuerbar macht, genügt es, den Fall zu betrachten, für den $\tilde{\mathbf{A}}_{21}$ und $\tilde{\mathbf{B}}_2$ skalare Nullen sind [49]. Mit Hilfe dieses Ansatzes kann μ_C durch die Minimierung der Norm $\|[\tilde{\mathbf{A}}_{21}, \tilde{\mathbf{B}}_2]\|$ berechnet werden. Dies führt, wie von WICKS und DeCARLO in [49] gezeigt, zu dem Optimierungsproblem

$$\mu_C(\mathbf{A}, \mathbf{B}) = \inf_{\mathbf{q} \in \mathbb{C}^n, \|\mathbf{q}\|=1} \{ \|[\mathbf{q}^H \mathbf{A}(\mathbf{E} - \mathbf{q}\mathbf{q}^H), \mathbf{q}^H \mathbf{B}]\| \}, \quad (3.23)$$

welches dem Optimierungsproblem aus Gleichung (3.21) entspricht. Die Zusammenhänge zwischen den Gleichungen (3.3) und (3.23) motivieren die in [49] vorgestellten Algorithmen. In Relation (3.23) sucht man nach einem Eigenvektor und in Relation (3.3) nach einem Eigenwert des nächstliegenden nicht steuerbaren Systems.

Es sei s^* der Wert der (3.3) minimiert und \mathbf{q}^* der Vektor, der zu einem Minimum von (3.23) führt. Dann gelten die Zusammenhänge

$$s^* = \mathbf{q}^{*H} \mathbf{A} \mathbf{q}^*$$

und

$$\mathbf{q}^{*H} \mathbf{A} \mathbf{v} = \sigma_{\min}([s^* \mathbf{E} - \mathbf{A}, \mathbf{B}]).$$

D.h. der Vektor \mathbf{q}^* ist linker Singulärvektor des kleinsten Singulärwertes von $[s^* \mathbf{E} - \mathbf{A}, \mathbf{B}]$. Die in [49] vorgestellten Algorithmen basieren auf dem „parallelen“ Lösen der beiden Optimierungsaufgaben (3.3) und (3.23). Es wird davon ausgegangen, dass, wenn ein Schätzwert \hat{s} gefunden wurde, der linke Singulärvektor des kleinsten Singulärwertes von $[\hat{s} \mathbf{E} - \mathbf{A}, \mathbf{B}]$ berechnet werden kann, welcher eine Approximation von \mathbf{q}^* darstellt. Das ermittelte $\hat{\mathbf{q}}$ kann wiederum für eine bessere Schätzung von s^* genutzt werden, indem $\hat{s} = \hat{\mathbf{q}}^H \mathbf{A} \hat{\mathbf{q}}$ berechnet wird (vgl. Abbildung 3.8).

In [49] werden Algorithmen vorgestellt, die diese Idee nutzen, aber die aufwendige Singulärwertzerlegung umgehen. Weiters wird gezeigt, dass die Algorithmen in ein lokales Minimum konvergieren, wodurch das Ergebnis stark vom Startwert abhängig ist. Auch hier kann auf die gleiche Weise wie bei der BFGS-Methode mit Hilfe von GUs Test entschieden werden, ob es sich bei dem gefunden Minimum um ein lokales oder um das globale Minimum handelt.

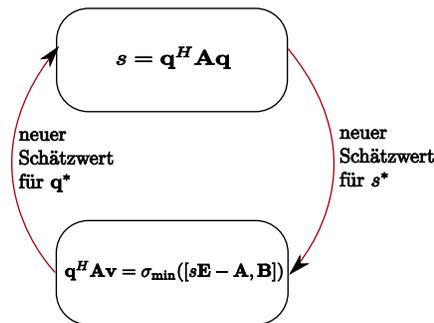


Abbildung 3.8: Algorithmus nach WICKS und DeCARLO

3.2.1.4 Trisection-Algorithmus mit GUs Test

Mit Hilfe des in [6] vorgestellten Algorithmus wird durch wiederholte Anwendung von GUs Test eine obere Schranke U und eine untere Schranke L von μ_C bestimmt. Die Arbeitsweise des Algorithmus soll anhand von Abbildung 3.9 verdeutlicht werden.

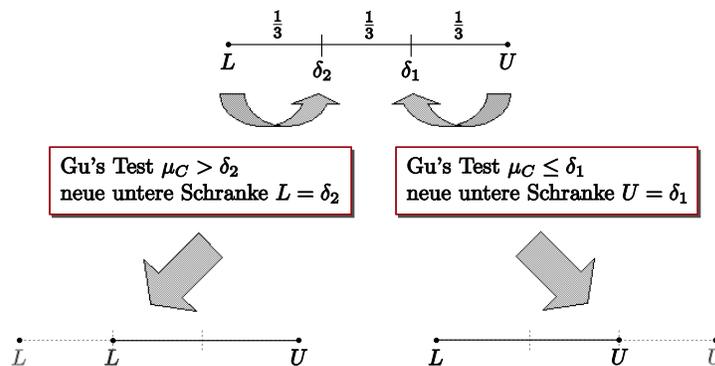


Abbildung 3.9: Trisection-Algorithmus

Je nach Ausgang von GUs Test wird entweder die obere oder die untere Schranke verschoben, so dass sich das Intervall $[L, U]$ bei jedem Durchlauf um ein Drittel reduziert. Dies geschieht so lange, bis der Abstand zwischen den beiden Schranken unterhalb eines vorgegebenen, positiven Toleranzwertes tol ist.

Algorithmus 3.2: Trisection

1. Setze $\delta_1 = L + 2(U - L)/3$ und $\delta_2 = L + (U - L)/3$.
2. Wenn $\mu_C(\mathbf{A}, \mathbf{B}) \leq \delta_1$, dann setze $U = \delta_1$;
andernfalls ist $\mu_C(\mathbf{A}, \mathbf{B}) > \delta_2$ und man setze $L = \delta_2$.
3. Wiederhole Schritt 2 bis $U - L < tol$.

Um die Genauigkeit bei der Bestimmung von μ_C zu erhöhen, kann tol nicht beliebig klein vorgegeben werden. Sind die beiden Werte δ_1 und δ_2 zu dicht aneinander, kann dies zu numerischen Schwierigkeiten führen, da der notwendige Vergleich der imaginären Eigenwerte von $\mathbf{H}(\alpha, \delta_1)$ und $\mathbf{H}(\alpha + \eta, \delta_1)$ auf Grund von Rundungsfehlern nicht mit beliebiger Zuverlässigkeit durchgeführt werden kann [6].

3.2.1.5 Minimumsuche nach GAO und NEUMANN

Bei dem von GAO und NEUMANN in [16] vorgestellten Algorithmus wird von folgendem Lemma ausgegangen:

Lemma 7: Es seien s_0, p komplexe Zahlen und $|p| = 1$. Für die Gerade

$$s = s_0 + kp, \quad k \in \mathbb{R}$$

gilt

$$\alpha \geq \alpha^*(s_0) := \min_{k \in \mathbb{R}} \{ \sigma_{\min}([(s_0 + kp)\mathbf{E} - \mathbf{A}, \mathbf{B}]) \}$$

genau dann, wenn die Matrix

$$\mathbf{H}(s_0, \alpha) = j \begin{pmatrix} \bar{p}(s_0\mathbf{E} - \mathbf{A}) & \bar{p}(\mathbf{B}\mathbf{B}^H - \alpha^2\mathbf{E}) \\ -p\mathbf{E} & p(\bar{s}_0\mathbf{E} - \mathbf{A}^H) \end{pmatrix} \quad (3.24)$$

einen rein imaginären Eigenwert besitzt.

Obiges Lemma geht von folgendem Ansatz aus: Der Wert α ist ein Singulärwert der Matrix $[s\mathbf{E} - \mathbf{A}, \mathbf{B}]$. Deshalb existieren die Vektoren $\mathbf{g}, \mathbf{h}, \mathbf{z}$, so dass

$$(s\mathbf{E} - \mathbf{A}, \mathbf{B}) \begin{pmatrix} \mathbf{g} \\ \mathbf{h} \end{pmatrix} = \alpha \mathbf{z} \quad \text{und} \quad \begin{pmatrix} s\mathbf{E} - \mathbf{A}^H \\ \mathbf{B}^H \end{pmatrix} \mathbf{z} = \alpha \begin{pmatrix} \mathbf{g} \\ \mathbf{h} \end{pmatrix} \quad (3.25)$$

mit $\mathbf{h} = \frac{1}{\alpha} \mathbf{B}^H \mathbf{z}$ gilt. Durch Variablensubstitution wird das Gleichungssystem in ein System niedrigerer Ordnung umgeformt. Man erhält dadurch

$$\mathbf{H}(s_0, \alpha) \mathbf{y} = -i k \mathbf{y} \quad \text{mit} \quad \mathbf{y} = \begin{pmatrix} \mathbf{g} \\ \frac{1}{\alpha} \mathbf{z} \end{pmatrix}$$

und somit Lemma 7.

Auch GUs Test benutzt das Gleichungssystem (3.25), welches durch geeignete Umformung auf ein System niedrigerer Ordnung reduziert wird. In beiden Fällen wird das ursprüngliche Problem in ein Eigenwertproblem umgeformt. Der Nachteil der von GAO und NEUMANN durchgeführten Umformung ist, dass bei der Eigenwertbestimmung von $\mathbf{H}(s_0, \alpha)$ ein Punkt s_0 in der komplexen Ebene benötigt wird. Deshalb muss z.B. ein geeignetes Gitter über die komplexe Ebene gespannt werden, an dessen Gitterpunkten das Eigenwertproblem ausgewertet wird, um das globale Minimum und somit μ_C zu bestimmen. Die Genauigkeit bei der Bestimmung des Distanzmaßes μ_C hängt daher von der Enge des Gitternetzes ab. Bei der Umformung nach GU ist die Auswertung des Eigenwertes aufwendiger, hat jedoch den Vorteil, dass für das Eigenwertproblem kein Punkt in der komplexen Ebene gewählt werden muss. Die von GAO und NEUMANN verwendete Umformung kann jedoch auch für GUs Test eingesetzt werden, um

die Berechnungsdauer des Tests zu reduzieren (siehe Abschnitt 3.2.6.1).

Eine Möglichkeit, wie der Ansatz von GAO und NEUMANN für die Bestimmung des Distanzmaßes μ_C genutzt werden kann bzw. wie das bereits angesprochene Gitternetz aufgespannt werden kann, wird in [16] vorgestellt und im Folgenden kurz zusammengefasst.

Für die in Lemma 7 angeführte Gerade $s = s_0 + kp$ wähle man die Parameter $p = i$ und $s_0 \in \mathbb{R}$. Damit entspricht s_0 dem Realteil und k dem Imaginärteil der komplexen Zahl s . Eine Vorgabe von s_0 bedeutet die Auswahl einer zur imaginären Achse parallelen Gerade. Für ein vorgegebenes α kann mittels des Lemmas entschieden werden, ob der Wert α größer gleich oder kleiner als das Minimum $\alpha^*(s_0)$ auf dieser Geraden ist. Um den gesamten Bereich (für eine mögliche Wahl siehe Abschnitt 3) der komplexen Ebene in dem das Minimum gesucht wird abzudecken, wird dieser in immer kleiner werdende Dreiecke unterteilt (siehe Abbildung 3.10).

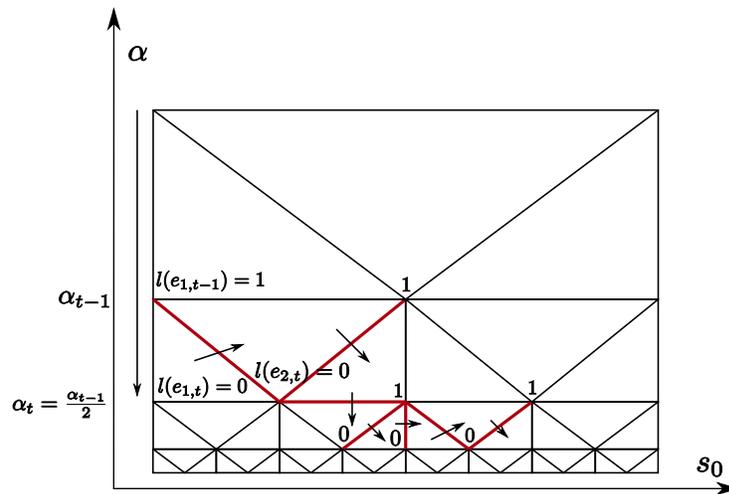


Abbildung 3.10: Vorgangswise des Algorithmus nach GAO und NEUMANN

Jedem Eckpunkt $e_{i,j}$ mit den Koordinaten $(s_{0,i}, \alpha_j)$ kann ein eindeutiger Wert

$$l(e_{i,j}) = \begin{cases} 0 & \mathbf{H}(s_{0,i}, \alpha_j) \text{ aus Gleichung (3.24) hat keinen rein imaginären} \\ & \text{Eigenwert, d.h. } \alpha < \alpha^*(s_0) \\ 1 & \mathbf{H}(s_i) \text{ hat einen rein imaginären Eigenwert,} \\ & \text{d.h. } \alpha \geq \alpha^*(s_0) \end{cases} \quad (3.26)$$

zugeordnet werden. Der Algorithmus bestimmt Schranken von μ_C , indem das durch die Dreiecke aufgespannte Gitter folgendermaßen durchlaufen wird:

Ausgangspunkt ist die linke obere Ecke $e_{1,1}$ des Gitters mit den Koordinaten $(s_{0,1}, \alpha_1)$. Da gilt $\alpha_1 \geq \mu_C$ wird dieser Ecke gemäß (3.26) der Wert $l(e_{1,1}) = 1$ zugeordnet. Durch Halbieren des Wertes α_i ($i = 1, 2, \dots$) wird das Gitter senkrecht nach unten durchlaufen und die dazugehörigen Werte $l(e_{1,i})$ bestimmt. Dies wird so lange wiederholt bis eine Ecke $e_{1,t}$ mit $l(e_{1,t}) = 0$ gefunden wird. Es gilt $\alpha_t < \alpha^*(s_0)$, d.h. der Wert α_t ist kleiner als das Minimum der durch s_0 verlaufenden Parallele zur imaginären Achse. Als nächstes muss somit der Wert von s_0 verändert werden, was einen Schritt in die waagrechte Richtung zur Ecke $e_{2,t}$ entspricht. Ist der Wert von $l(e_{2,t})$ gleich 1, dann ist $[e_{1,t}, e_{2,t}]$ eine sogenannte DV-Kante (different-valued), d.h. die Eckpunkte haben nicht den gleichen Wert. Ist $l(e_{2,t}) = 0$, muss $[e_{1,t-1}, e_{2,t}]$ eine DV-Kante sein.

Ausgehend von der gefundenen DV-Kante gelangt man in das Dreieck $\langle e_{2,t}, e_{1,t-1}, e_{2,t-1} \rangle$, dessen Eckpunkt $e_{2,t-1}$ als einzigem noch kein Wert zugeordnet wurde. Durch die Bestimmung von $l(e_{2,t-1})$ erhält man eine weitere DV-Kante, die in das nächste Dreieck führt, bis entweder der obere oder der rechte Rand des zu durchsuchenden Bereichs erreicht wird. Der Algorithmus bestimmt dann das Minimum der Koordinate α_i aller Eckpunkte $e_{i,j}$ mit $l(e_{i,j}) = 1$, die durchlaufen wurden. Aus dem so erhaltenen Minimum werden schlussendlich die obere und die untere Schranke von μ_C berechnet.

In Abbildung 3.10 ist ein Teil des Gitters mit der Vorgehensweise des Algorithmus für einige Schritte dargestellt. Die roten Linien entsprechen den DV-Kanten, die in das nächste Dreieck führen.

Die Arbeitsweise des Algorithmus wird in Abschnitt 4.1.1 an einem Beispiel erläutert.

3.2.1.6 Berechnung mittels KRONECKER Form

Das von BOLEY in [4] vorgestellte Verfahren zur Bestimmung von μ_C nutzt Eigenschaften der KRONECKER kanonischen Form (KCF), in die jede Matrix der Form $\mathbf{F} + s\mathbf{G}$ (sogenannte *Matrizenbüschel*) überführt werden kann. Die KRONECKER kanonische Form ist eine Blockdiagonalmatrix, bestehend aus den Blöcken \mathbf{N} , \mathbf{L} , \mathbf{L}^T und/oder $\mathbf{J} + s\mathbf{E}$ (siehe Anhang A.3). Die Theorie der Matrizenbüschel bzw. der KRONECKER kanonischen Form kann für die Überprüfung der Steuerbarkeit des Systems (\mathbf{A}, \mathbf{B}) angewendet werden. Hierfür wird das Büschel

$$\mathbf{F} + s\mathbf{G} := \begin{bmatrix} \mathbf{A}^T \\ \mathbf{B}^T \end{bmatrix} + s \begin{bmatrix} -\mathbf{E} \\ \mathbf{0} \end{bmatrix}$$

eingeführt. Nach dem HAUTUS-Kriterium ist ein LZI-System genau dann steuerbar, wenn dieses Büschel für jeden beliebigen komplexen Wert s den vollen Rang besitzt. Auf Grund der speziellen Struktur des hier betrachteten Büschels kann die KCF nur \mathbf{L}^T und/oder $\mathbf{J} + s\mathbf{E}$ Blöcke besitzen. Dies entspricht dem sogenannten singulären und regulären Anteil des Büschels. Ist das System (\mathbf{A}, \mathbf{B}) nicht steuerbar, verliert das Büschel $\mathbf{F} + s\mathbf{G}$ für einen endlichen Wert s den vollen Rang. Daraus folgt, dass ein regulärer Anteil der Form $\mathbf{J} + s\mathbf{E}$ existiert, denn es gibt keinen Wert s , der einen Rangverlust in einem \mathbf{L}^T Block verursachen kann.

Die Steuerbarkeit des Systems kann auch durch folgendes Eigenwertproblem beurteilt werden: Betrachtet man beliebige Matrizen \mathbf{C} und \mathbf{D} passender Dimension, erhält man das generalisierte Eigenwertproblem

$$[\mathbf{F}, \mathbf{C}]_{\mathbf{x}_R} = s[-\mathbf{G}, -\mathbf{D}]_{\mathbf{x}_R} \text{ bzw. } [\mathbf{F} + s\mathbf{G}, \mathbf{C} + s\mathbf{D}]_{\mathbf{x}_R} = \mathbf{0}. \quad (3.27)$$

Der Vektor \mathbf{x}_R wird zerlegt in

$$\mathbf{x}_R^T = [\mathbf{x}^T, \mathbf{y}^T]$$

mit \mathbf{x} , einem Vektor der Länge n , und \mathbf{y} , einem Vektor der Länge m . Für beliebige Matrizen \mathbf{C} und \mathbf{D} mit vollem Rang sind die folgenden Aussagen äquivalent:

- $\mathbf{F} + s\mathbf{G}$ ist *nicht* steuerbar.
- Es existiert ein von Null verschiedener Vektor \mathbf{x} und eine skalare Größe s_0 , so dass

$$\mathbf{F}\mathbf{x} = -s_0\mathbf{G}\mathbf{x},$$

d.h., für den Wert s_0 verliert die Matrix $\mathbf{F} + s\mathbf{G}$ den vollen Rang. Die KCF von $\mathbf{F} + s\mathbf{G}$ besitzt somit einen regulären Anteil, und der Wert s_0 ist der Eigenwert des nicht steuerbaren Teilsystems.

- Der zu s_0 gehörende Vektor \mathbf{y} des generalisierten Eigenwertproblems (3.27) ist gleich dem Nullvektor.

Bevor erläutert wird, wie diese Zusammenhänge zur Ermittlung von Schranken für das Distanzmaß μ_C genutzt werden, werden diese Beziehungen anhand des bereits in Beispiel 2.2 verwendeten Systems verdeutlicht.

Beispiel 3.1: Für das System

$$\mathbf{A} = \begin{bmatrix} \alpha & 1 \\ 0 & -2 \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

mit $n = 2$ und $m = 1$ ergibt sich für

$$\mathbf{F} = \begin{bmatrix} \alpha & 0 \\ 1 & -2 \\ 1 & 1 \end{bmatrix} \quad \text{und} \quad \mathbf{G} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \\ 0 & 0 \end{bmatrix}.$$

Daraus folgt

$$\mathbf{F} + s\mathbf{G} = \begin{bmatrix} \alpha - s & 0 \\ 1 & -2 - s \\ 1 & 1 \end{bmatrix}.$$

Es werden nun für zwei verschiedene Werte von α die KCF sowie die Eigenwerte und Eigenvektoren des generalisierten Eigenwertproblems (3.27) berechnet.

Steuerbares System: $\alpha = -1$.

Der Index $k = 2$ ist der kleinstmögliche Wert, der die Ungleichung aus Lemma 15 (siehe Anhang A.3) erfüllt. Daraus folgt, dass für das hier untersuchte System die Gleichung $(\mathbf{F}^T + s\mathbf{G}^T)\mathbf{x} = \mathbf{0}$ ein nichttriviales Lösungspolynom $\mathbf{x}(s)$ zweiten Grades besitzt. Somit gilt $\eta_1 = 2$ und die KCF besteht aus einer Matrix \mathbf{L}^T der Dimension $(3, 2)$. Da dies der Dimension des betrachteten Bündels $\mathbf{F} + s\mathbf{G}$ entspricht, folgt daraus, dass die KCF nur aus diesem einen Block besteht. Damit erhält man

$$\mathbf{U}^{-1}(\mathbf{F} + s\mathbf{G})\mathbf{V} = \mathbf{S} + s\mathbf{T} = \begin{bmatrix} s & 0 \\ 1 & s \\ 0 & 1 \end{bmatrix}$$

mit den Transformationsmatrizen

$$\mathbf{U}^{-1} = \begin{bmatrix} 3 & 1 & 2 \\ -1 & -1 & -3 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{und} \quad \mathbf{V} = \frac{1}{2} \begin{bmatrix} -1 & -1 \\ 1 & 3 \end{bmatrix}.$$

Da die KCF keinen regulären Anteil besitzt, ist das System steuerbar.

Bei der Lösung des generalisierten Eigenwertproblems ergeben sich für die Eigenwerte bzw. die Eigenvektoren folgende Werte:

$$\begin{aligned} s_1 &= 1,73 \\ \mathbf{x}_{R,1}^T &= [0,37 \quad 0,37 \quad 1] \\ s_2 &= -1,73 \\ \mathbf{x}_{R,2}^T &= [-1 \quad -1 \quad 0,37] \\ s_3 &= -2 \\ \mathbf{x}_{R,3}^T &= [-0,5 \quad -1 \quad 0,5] \end{aligned}$$

Der Wert y (entspricht der dritten Komponente des Vektors \mathbf{x}_R) ist für alle Eigenvektoren ungleich Null, da das System steuerbar ist.

Nicht steuerbares System: $\alpha = -3$.

Die Ungleichung aus Lemma 15 ist für den kleinstmöglichen Wert $k = 1$ erfüllt, weshalb für diesen Fall Gleichung $(\mathbf{F}^T + s\mathbf{G}^T)\mathbf{x} = \mathbf{0}$ ein nichttriviales Lösungspolynom $\mathbf{x}(s)$ ersten Grades besitzt. Somit gilt $\eta_1 = 1$, und die KCF besitzt einen \mathbf{L}^T Block der Dimension $(2, 1)$ und einen regulären Anteil der Dimension $(1, 1)$. Mit den Matrizen

$$\mathbf{U}^{-1} = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 0 & -1 \\ -1 & 2 & -2 \end{bmatrix} \quad \text{und} \quad \mathbf{V} = \frac{1}{3} \begin{bmatrix} -2 & 1 \\ -1 & -1 \end{bmatrix}$$

erhält man die KCF

$$\mathbf{S} + s\mathbf{T} = \begin{bmatrix} \boxed{s} & 0 \\ 1 & 0 \\ 0 & \boxed{3+s} \end{bmatrix}.$$

Da die Matrix $\mathbf{S} + s\mathbf{T}$ für $s = -3$ den vollen Rang verliert, gilt $s_0 = -3$. Dieses s_0 entspricht genau dem Eigenwert des nicht steuerbaren Teilsystems.

Betrachtet man nun die Eigenwerte und Eigenvektoren des generalisierten Eigenwertproblems, ergeben sich folgende Werte:

$$\begin{aligned} s_1 &= 1,56 \\ \mathbf{x}_{R,1}^T &= [0,22 \quad 0,34 \quad 1] \\ s_2 &= -3 \\ \mathbf{x}_{R,2}^T &= [-1 \quad 1 \quad 0] \\ s_3 &= -2,56 \\ \mathbf{x}_{R,3}^T &= [0,39 \quad -1 \quad 0,17]. \end{aligned}$$

Man sieht, dass für den zum Eigenwert $s_2 = -3$ gehörenden Eigenvektor $y = 0$ gilt, da das System nicht steuerbar ist.

3.2.1.6.1 Ermittlung von oberen Schranken für μ_C Für die Bestimmung einer oberen Schranke von μ_C wird eine Störung $\delta\mathbf{F}$ betrachtet, so dass $\mathbf{F} + \delta\mathbf{F} + s\mathbf{G}$ einem nicht steuerbaren System

entspricht. Besitzt \mathbf{G} den vollen Rang und ist das System $\mathbf{F} + \delta\mathbf{F}$ nicht steuerbar, dann muss die KRONECKER Form der Matrix $\mathbf{F} + \delta\mathbf{F} + s\mathbf{G}$ einen regulären Teil haben oder es existiert in Äquivalenz hierzu ein Eigenwert s_i , für den $\mathbf{y}_i = \mathbf{0}$ gilt. Durch Umformen von Gleichung (3.27) zu

$$-\underbrace{(s_i\mathbf{G} + \mathbf{F})\mathbf{x}_i}_{=: \mathbf{r}_i} = (\mathbf{C} + s_i\mathbf{D})\mathbf{y}_i$$

sieht man, dass die Wahl der Störung

$$\delta\mathbf{F}_i = -\frac{\mathbf{r}_i\mathbf{x}_i^T}{\|\mathbf{x}_i\|^2} = -(s_i\mathbf{G} + \mathbf{F})\frac{\mathbf{x}_i\mathbf{x}_i^T}{\|\mathbf{x}_i\|^2} = (\mathbf{C} + s_i\mathbf{D})\frac{\mathbf{y}_i}{\|\mathbf{x}_i\|}\frac{\mathbf{x}_i^T}{\|\mathbf{x}_i\|}$$

für jedes i , für das gilt $1 \leq i \leq m + n$, zu einem nicht steuerbaren System $\mathbf{F} + \delta\mathbf{F}_i + s_i\mathbf{G}$ führt. Eine obere Schranke für μ_C ist somit gegeben durch

$$\beta_1 = \min_i \{\|\delta\mathbf{F}_i\|\} = \min_i \left\{ \frac{\|(\mathbf{C} + s_i\mathbf{D})\mathbf{y}_i\|}{\|\mathbf{x}_i\|} \right\}.$$

BOLEY schlägt basierend auf der Schranke β_1 eine weitere obere Schranke vor: Es seien σ_i , \mathbf{u}_i und \mathbf{v}_i der Singulärwert und zugehörige Singulärvektoren von $\mathbf{F} + s_i\mathbf{G}$ für den in obiger Gleichung gefundenen Wert i . Dann ist

$$\delta\mathbf{F}'_i := -\sigma_i\mathbf{u}_i\mathbf{v}_i^H,$$

eine in Bezug zu $\delta\mathbf{F}_i$ kleinere Störung, die zu einem nicht steuerbaren System führt. Dadurch erhalten wir eine kleinere obere Schranke

$$\beta_2 = \min_i \{\sigma_{\min}(\mathbf{F} + s_i\mathbf{G})\}.$$

Für die Berechnung der oberen Schranken von μ_C wird durch die spezielle Wahl von

$$\mathbf{D} = [\mathbf{0}, -\mathbf{E}]^T$$

Gleichung (3.27) zu einem gewöhnlichen Eigenwertproblem. Um das Eigenwertproblem so gut wie möglich zu konditionieren, wird die Matrix \mathbf{C} so gewählt, dass sie orthonormale Spalten besitzt.

3.2.1.6.2 Ermittlung einer unteren Schranke für μ_C In [4] bestimmt BOLEY mit Hilfe der Matrix-Störungstheorie folgendermaßen eine untere Schranke für μ_C . Wir betrachten Gleichung (3.3)

$$\mu(\mathbf{A}, \mathbf{B}) = \inf_{s \in \mathbb{C}} \{\sigma_{\min}([s\mathbf{E} - \mathbf{A}, \mathbf{B}])\} = \inf_{s \in \mathbb{C}} \{\sigma_{\min}(\mathbf{F} + s\mathbf{G})\}.$$

Es sei s^* der komplexe Wert, der obige Relation minimiert. Mit $\sigma(s)$ wird der kleinste Singulärwert der Matrix $\mathbf{F} + s\mathbf{G}$ bezeichnet. Auf Grund des Theorems nach MIRSKY (siehe Anhang A.4.2) gilt

$$|\sigma(\lambda_i) - \sigma(s^*)| \leq \|(\mathbf{F} + \lambda_i\mathbf{G}) - (\mathbf{F} + s^*\mathbf{G})\|.$$

Eine weitere Ungleichung kann mit Hilfe des BAUER-FIKE-Theorems (siehe Anhang A.4.1) ermittelt werden. BOLEY nutzt folgende leicht modifizierte Variante dieses Theorems:

Theorem 8 (BAUER-FIKE (modifiziert)): Gegeben sei eine (n, n) -Matrix \mathbf{A} mit den einfachen Eigenwerten $\lambda_1, \dots, \lambda_n$. Es sei

$$\mathbf{X}_R := (\mathbf{x}_{R,1}, \dots, \mathbf{x}_{R,n})$$

die Matrix von zugehörigen linear unabhängigen Eigenvektoren $\mathbf{x}_{R,i}$. Es seien $\mathbf{\Delta}$ eine beliebige (n, n) -Matrix und $\tilde{\lambda}$ ein Eigenwert der Matrix $\mathbf{A} + \mathbf{\Delta}$. Dann gilt für mindestens ein λ_i , für das gilt

$$|\tilde{\lambda} - \lambda_i| \leq K_i \|\mathbf{\Delta}\| \quad 1 \leq i \leq n$$

mit

$$K_i = \min \left\{ K(\mathbf{X}_R) = \|\mathbf{X}_R\| \|\mathbf{X}_R^{-1}\|, n \frac{\|\mathbf{x}_{L,i}\| \|\mathbf{x}_{R,i}\|}{\mathbf{x}_{L,i}^H \mathbf{x}_{R,i}} \right\}.$$

Dieses Theorem wird nun auf das erweiterte Matrizenbüschel nach Gleichung (3.27)

$$([\mathbf{F}, \mathbf{C}] + s[\mathbf{G}, \mathbf{D}])\mathbf{x}_R = \mathbf{0}$$

mit

$$\mathbf{F} = \begin{bmatrix} \mathbf{A}^T \\ \mathbf{B}^T \end{bmatrix} \quad \text{und} \quad [\mathbf{G}, \mathbf{D}] = -\mathbf{E}$$

angewendet. Obige Gleichung beschreibt somit ein gewöhnliches Eigenwertproblem der quadratischen Matrix $[\mathbf{F}, \mathbf{C}]$. Es sei τ der kleinste Singulärwert von $[\mathbf{F}, \mathbf{C}] - s^* \mathbf{E}$. Dann gilt $\tau \leq \sigma(s^*)$, da das Hinzufügen von Spalten zu einer Matrix mit mehr Zeilen als Spalten den kleinsten Singulärwert nur verkleinern kann. Es kann eine Matrix $\mathbf{\Delta}$ mit $\|\mathbf{\Delta}\| = 1$ gefunden werden, so dass der Wert s^* ein Eigenwert von $[\mathbf{F}, \mathbf{C}] + \tau \mathbf{\Delta}$ ist. Für mindestens einen der Eigenwerte λ_i ($i = 1, \dots, n$) von $[\mathbf{F}, \mathbf{C}]$ impliziert damit das modifizierte BAUER-FIKE-Theorem, dass

$$|\lambda_i - s^*| \leq \tau K_i$$

gilt. Somit erhält man die Ungleichungen

$$|\sigma(\lambda_i) - \sigma(s^*)| \leq \|(\mathbf{F} + \lambda_i \mathbf{G}) - (\mathbf{F} + s^* \mathbf{G})\| \leq |\lambda_i - s^*| \leq \tau K_i \leq \sigma(s^*) K_i.$$

Daraus folgt $\sigma(\lambda_i) \leq \sigma(s^*) (K_i + 1)$, wodurch man die untere Schranke

$$\mu_C = \sigma(s^*) \geq \frac{\sigma(\lambda_i)}{K_i + 1} \geq \frac{\beta_2}{K_i + 1}$$

erhält.

3.2.1.7 Bestimmung von Schranken durch Eigenwertabschätzung

Die in [20] von HE vorgestellte Abschätzung basiert auf folgendem Theorem:

Theorem 9 (Eigenwertabschätzung): Es sei \mathbf{M} eine (n, n) -Diagonalmatrix

$$\mathbf{M} = \text{diag}(d_i) \quad i = 1, \dots, n$$

mit $d_1 < d_2 < \dots < d_n$. Wird die Matrix \mathbf{M} durch eine additive Störung $\mathbf{z}\mathbf{z}^T$ verändert

$$\tilde{\mathbf{M}} := \mathbf{M} + \mathbf{z}\mathbf{z}^T \quad \text{mit} \quad \mathbf{z} = [z_1, z_2, \dots, z_n]^T,$$

dann gilt für den kleinsten Eigenwert der so entstandenen Matrix $\lambda_1 = \min\{\Lambda(\tilde{\mathbf{M}})\}$

$$\lambda_1 \geq \min \left\{ d_1 + \frac{1}{2}(d_2 - d_1), d_1 + \frac{1}{2}z_1^2, \frac{(d_2 - d_1)z_1^2}{4l^2} + d_1 \right\} \quad \text{mit } l = \sum_{j=2}^n z_j^2.$$

Obiges Theorem kann für die Ermittlung einer unteren Schranke für das Distanzmaß μ_C benutzt werden. Für das in Modalform vorliegende System

$$\frac{dz}{dt} = \mathbf{X}_L \mathbf{A} \mathbf{X}_R z + \mathbf{X}_L \mathbf{B} u =: \hat{\mathbf{A}} z + \hat{\mathbf{B}} u$$

mit

$$\mathbf{X}_L \mathbf{A} \mathbf{X}_R = \hat{\mathbf{A}} = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{bmatrix} \quad \text{und} \quad \mathbf{X}_L \mathbf{B} = \hat{\mathbf{B}} = \begin{bmatrix} \hat{\mathbf{b}}_1^T \\ \hat{\mathbf{b}}_2^T \\ \vdots \\ \hat{\mathbf{b}}_n^T \end{bmatrix}$$

ist μ_C^2 der kleinste Eigenwert der Matrix

$$(s^* \mathbf{E} - \hat{\mathbf{A}})(s^* \mathbf{E} - \hat{\mathbf{A}})^H + \hat{\mathbf{B}} \hat{\mathbf{B}}^H$$

mit s^* dem extremalen Wert von (3.3). Obige Matrix besitzt die gleiche Struktur wie $\tilde{\mathbf{M}}$, wodurch Theorem 9 eine untere Schranke für das Distanzmaß $\mu_C(\hat{\mathbf{A}}, \hat{\mathbf{B}})$ angibt. Mittels der Beziehung

$$\frac{1}{K(\mathbf{X}_R)} \mu_C(\hat{\mathbf{A}}, \hat{\mathbf{B}}) \leq \mu_C(\mathbf{A}, \mathbf{B})$$

mit der Konditionszahl $K(\mathbf{X}_R) = \|\mathbf{X}_R\| \|\mathbf{X}_R^{-1}\|$ erhält man folgendes Theorem zur Bestimmung einer unteren Schranke für das Distanzmaß μ_C des Systems (\mathbf{A}, \mathbf{B}) :

Theorem 10: Wir definieren folgende Größen

$$\begin{aligned} \gamma_i &= \max_j \frac{|\hat{b}_{ij}|}{2\sqrt{\sum_{k, k \neq i} |\hat{b}_{kj}|^2}}, \\ \vartheta &= \min_{i, \gamma_i < \sqrt{0,5}} \left(\frac{\gamma_i}{\sqrt{2}} \min_{k, k \neq i} |\lambda_i - \lambda_k| \right), \\ \hat{\beta} &= \max_j \min_i |\hat{b}_{ij}| \quad \text{und} \\ \hat{\lambda} &= \min_{i, j} |\lambda_i - \lambda_j| \quad \text{mit } i \neq j. \end{aligned}$$

Für

$$\|\mathbf{X}_R\| = 1$$

gilt

$$\frac{1}{K(\mathbf{X}_R)} \min \left\{ \frac{\hat{\beta}}{\sqrt{2}}, \frac{\hat{\lambda}}{2}, \vartheta \right\} \leq \mu_C.$$

Als obere Schranke schlägt HE Folgendes vor: Für die Indizes i_1 und j_1 gelte

$$|\lambda_{i_1} - \lambda_{j_1}| = \hat{\lambda}.$$

Es seien i_2, j_2 die Indizes, so dass die Gleichung

$$\hat{\beta} = |\hat{b}_{i_2, j_2}|$$

erfüllt ist. Dann gilt

$$\mu_C \leq \min\{\sigma(\lambda_{i_1}), \sigma(\lambda_{j_1}), \sigma(\lambda_{i_2})\}.$$

Hierbei entspricht $\sigma(s)$ dem kleinsten Singulärwert der Matrix $[s\mathbf{E} - \mathbf{A}, \mathbf{B}]$. Diese obere Schranke für das Distanzmaß μ_C ist durch Theorem 10 motiviert. Aus diesem Theorem folgt, dass $\sigma(\lambda)$ für die Werte $\lambda_{i_1}, \lambda_{j_1}$ und λ_{i_2} nahe bei der unteren Schranke liegen könnte und somit eine gute obere Schranke darstellt [20].

Weiters gilt

$$\mu_C \leq \min_k \{\sigma(\lambda_k)\} =: \mu_\lambda \leq \min\{\sigma(\lambda_{i_1}), \sigma(\lambda_{j_1}), \sigma(\lambda_{i_2})\} \quad \text{mit } k = 1, \dots, n,$$

d.h., der kleinste Wert von σ , ausgewertet für alle Eigenwerte der Matrix \mathbf{A} , liefert eine kleinere obere Schranke für μ_C als die von HE vorgeschlagene Abschätzung. Für die Berechnung der Schranke μ_λ sind bei verschiedenen Eigenwerten n Singulärwertzerlegungen durchzuführen, während für die Bestimmung der Schranke nach HE unabhängig von der Systemordnung maximal drei Singulärwertzerlegungen benötigt werden. Allerdings muss für die Bestimmung dieser Schranke das System in Modalform transformiert werden. Auch für die Bestimmung der unteren Schranke wird die Modalform benötigt. Die von HE vorgeschlagenen Schranken können daher nur berechnet werden, wenn die Modalform existiert, wenn also die Matrix \mathbf{A} linear unabhängige Eigenvektoren besitzt.

3.2.2 Numerische Bestimmung der Größe μ_R

In den folgenden Abschnitten werden Ansätze für die Berechnung von μ_R vorgestellt, die für die Störungen $\delta\mathbf{A}$ und $\delta\mathbf{B}$ in (3.2) nur reelle Werte zulassen. Durch diese Voraussetzung ist s^* eine reelle Größe. Somit kann die Suche nach dem Minimum auf die reelle Achse beschränkt werden. Die Lage der lokalen Minima in der komplexen Ebene sowie deren Werte ändern sich stetig, wenn \mathbf{A} oder \mathbf{B} sich stetig ändern. Da bei der Bestimmung von μ_R nur die reelle Achse betrachtet wird, kann es zu Unstetigkeiten kommen, wenn das globale Minimum nicht auf der reellen Achse liegt (d.h. $\mu_C \neq \mu_R$).

Verfolgt man, wie schon in Abschnitt 3.2.1.3, die Idee, dass gemäß der KALMAN-Zerlegung für ein nicht steuerbares System die Matrizen $\tilde{\mathbf{A}}_{21}$ und $\tilde{\mathbf{B}}_2$ in Gleichung (3.22) Null sind, so kann auch für die Berechnung von μ_R das Minimum von $\|[\tilde{\mathbf{A}}_{21}, \tilde{\mathbf{B}}_2]\|$ herangezogen werden. Wie in [49] gezeigt, muss im reellen Fall nicht nur eine Matrix $\tilde{\mathbf{A}}_{21}$ mit Rang eins sondern auch mit Rang zwei berücksichtigt werden. Diese Minimierung führt nach WICKS und DeCARLO in [49] zu dem Optimierungsproblem

$$\mu_R(\mathbf{A}, \mathbf{B}) = \min_{\mathbf{Q} \in \{\mathbb{R}^{n \times 1} \cup \mathbb{R}^{n \times 2}\}, \|\mathbf{Q}\|=1} \{ \|[\mathbf{Q}^T \mathbf{A}(\mathbf{E} - \mathbf{Q}\mathbf{Q}^T), \mathbf{Q}^T \mathbf{B}]\| \}.$$

Das Optimierungsproblem über $\mathbf{Q} \in \mathbb{R}^{n \times 1}$ ist jenem aus Gleichung (3.23) ähnlich mit dem einzigen Unterschied, dass in diesem Fall nur reelle Werte für \mathbf{Q} zulässig sind. Die für die Berechnung von μ_R notwendige Optimierung über $\mathbf{Q} \in \mathbb{R}^{n \times 2}$ stellt ein deutlich aufwendigeres Optimierungsproblem dar.

Im Gegensatz dazu vereinfachen sich die anderen auf dem Optimierungsproblem basierenden Algorithmen zur Bestimmung von μ_C durch die auf die reelle Achse beschränkte Suche. Daher wird hier auf den Algorithmus nach WICKS und DeCARLO für reelle Störungen nicht näher eingegangen.

In den Abschnitten 3.2.2.1 und 3.2.2.2 werden zwei Methoden zur Mimumsuche auf der reellen Achse vorgestellt. Auch im reellen Fall werden in der Literatur Methoden zur Berechnung von Schranken für μ_R vorgestellt, die nicht auf dem Optimierungsproblem basieren. Diese werden in den Abschnitten 3.2.2.3 - 3.2.2.5 erläutert.

3.2.2.1 Berechnung von μ_R durch Nullstellensuche

Die Methoden zur Bestimmung von μ_C , welche nach lokalen Minima suchen, können für die Suche auf der reellen Achse genutzt werden. Die Suche ist beendet, wenn z.B. durch die reelle Variante von GÜs Test oder durch die Methode nach GAO und NEUMANN festgestellt wird, dass es sich bei dem gefundenen Minimum um das Globale handelt. Mit Hilfe der bereits in Abschnitt 3 definierten Funktion

$$f(s) = \mathbf{v}_n^H(s) \begin{bmatrix} \mathbf{u}_n(s) \\ 0 \end{bmatrix}$$

ergibt sich folgende einfache Bestimmung von lokalen Minima. Für die Ableitung von

$$\sigma(s) = \sigma_{\min}([s\mathbf{E} - \mathbf{A}, \mathbf{B}])$$

gilt mit $s = x + iy$

$$\frac{\partial \sigma(s)}{\partial x} = -\operatorname{Re} \{f(s)\} \quad \text{und} \quad \frac{\partial \sigma(s)}{\partial y} = -\operatorname{Im} \{f(s)\}.$$

Daraus folgt, dass die Nullstellen von $f(s)$ den Extremalen von $\sigma(s)$ entsprechen. Für die Bestimmung von μ_R müssen somit die reellen Nullstellen von $f(s)$ bestimmt werden. Hierfür kann beispielsweise der in [7] angeführte Algorithmus verwendet werden.

Algorithmus 3.3: Nullstellensuche

1. Finde ein Intervall $[a, b]$, sodass $f(a)f(b) < 0$.
2. Berechne $c = (a + b)/2$.
3. Ist $f(c)f(b) < 0$, dann setze $a = c$ und gehe zu Schritt 2.
Ist $f(c)f(a) < 0$, dann setze $b = c$ und gehe zu Schritt 2.
4. Wiederhole Schritt 2 und 3 bis $(b - a) < \text{tol}$.

Durch Bestimmen aller Nullstellen von $f(s)$ können alle lokalen Minima auf der reellen Achse bestimmt werden.

3.2.2.2 Minimumsuche nach GAO und NEUMANN

Für die direkte Suche nach dem globalen Minimum auf der reellen Achse können die bereits für die Bestimmung von μ_C vorgestellten Algorithmen in vereinfachter Form benutzt werden. Zum Beispiel: Basierend auf Lemma 7 wird in [16] ein Bisection Algorithmus vorgestellt, der einen Wert für μ_R mit einer vorgebbaren Toleranzschranke tol bestimmt. Hierfür wird in Gleichung (3.24) $s_0 = 0$ und $p = 1$ gesetzt. Es sei eine positive Schranke tol vorgegeben. Man wähle zwei positive Werte a und b , so dass

$$a \leq \mu_R \leq b.$$

Algorithmus 3.4: Bisection

1. Setze $c = (a + b)/2$ und berechne die Eigenwerte von $\mathbf{H}(c)$ aus Gleichung (3.24).
2. Hat $\mathbf{H}(c)$ einen rein imaginären Eigenwert, dann setze $b = c$, andernfalls setze $a = c$.
3. Falls $(b - a) \geq tol$ gehe zu Schritt 1, ansonsten ist der Algorithmus beendet.

Der beim letzten Schritt ermittelte Wert c ist eine Schätzung für μ_R .

Für die in dieser Arbeit durchgeführten Beispiele wurden gemäß Abschnitt 3 $a = 0$ und $b = \|\mathbf{A}\|_2$ gewählt.

Natürlich kann hier stattdessen der Trisection-Algorithmus (siehe Abschnitt 3.2.1.4) benutzt werden. An Stelle der Überprüfung nach GAO und NEUMANN kann auch eine reelle Variante von GUs Test verwendet werden.

3.2.2.3 Abschätzung von μ_R , falls μ_C vorliegt

Eine einfach zu berechnende Schranke für μ_R kann ermittelt werden, wenn μ_C bereits bekannt ist. GAHINET hat gezeigt, dass zwischen μ_C und μ_R der Zusammenhang

$$\mu_C \leq \mu_R \leq 2 \max \left\{ \sqrt{2} \mu_C, \|\mathbf{A} - \mathbf{A}^T\|^{1/3} \mu_C^{2/3} \right\} \quad (3.28)$$

besteht [14].

3.2.2.4 Bestimmung einer oberen Schranke mittels Quasi-KALMAN-Zerlegung

Eine Möglichkeit zur Bestimmung einer oberen Schranke basiert auf der in Anhang A.2 beschriebenen KALMAN-Zerlegung.

Lemma 11 (Quasi-KALMAN-Zerlegung [5]): Für ein steuerbares System (\mathbf{A}, \mathbf{B}) existiert eine orthogonale Matrix \mathbf{P} , so dass

$$\tilde{\mathbf{A}} = \mathbf{PAP}^T = \begin{bmatrix} \tilde{\mathbf{A}}_{11} & \tilde{\mathbf{A}}_{12} \\ \tilde{\mathbf{A}}_{21} & \tilde{\mathbf{A}}_{22} \end{bmatrix} \quad \tilde{\mathbf{B}} = \mathbf{PB} = \begin{bmatrix} \tilde{\mathbf{B}}_1 \\ \tilde{\mathbf{B}}_2 \end{bmatrix}$$

mit $\tilde{\mathbf{A}}_{11} \in \mathbb{R}^{r \times r}$, $\tilde{\mathbf{B}}_1 \in \mathbb{R}^{r \times m}$. Für beliebige r gilt

$$\|\tilde{\mathbf{A}}_{21}\| \leq \frac{\sigma_{r+1}}{\sigma_r} \|\mathbf{A}_B\| \text{ und } \|\tilde{\mathbf{B}}_2\| \leq \sigma_{r+1}.$$

Hierbei symbolisieren σ_i die Singulärwerte der Steuerbarkeitsmatrix \mathbf{S}_u und \mathbf{A}_B die Begleitmatrix von \mathbf{A} .

Ausgehend von diesem Theorem schlagen BOLEY und LU in [5] Folgendes vor: Zur Bestimmung einer oberen Schranke von μ_R werden nur diejenigen Störungen betrachtet die $\tilde{\mathbf{A}}_{21}$ und $\tilde{\mathbf{B}}_2$ zu skalaren Nullen und somit das System nicht steuerbar machen. Dadurch erhält man mit $r = n - 1$ die Schranke

$$\mu_R(\mathbf{A}, \mathbf{B}) \leq \left(1 + \frac{\|\mathbf{A}_B\|}{\sigma_{n-1}}\right) \sigma_n.$$

Theoretisch wäre aber auch eine Störung denkbar, die $\tilde{\mathbf{A}}_{21}$ und $\tilde{\mathbf{B}}_2$ zu einer Nullmatrix bzw. einen Nullvektor macht und eine kleinere Norm besitzt als jene Störung, die dazu führt, dass $\tilde{\mathbf{A}}_{21}$ und $\tilde{\mathbf{B}}_2$ zu skalaren Nullen werden. Deshalb liefert

$$\mu_R(\mathbf{A}, \mathbf{B}) \leq \min_r \left\{ \left(1 + \frac{\|\mathbf{A}_B\|}{\sigma_{r-1}}\right) \sigma_r \right\} \quad r = 1, \dots, n-1$$

eine engere obere Schranke.

Als einfache Abschätzung der Größenordnung von μ_C schlagen BOLEY und LU den größten Abstand zwischen zwei benachbarten Singulärwerten vor. Diese Abschätzung führt jedoch zu keinen zufriedenstellenden Ergebnissen, da sie in vielen Fällen deutlich zu groß ist bzw. auch falsche Schlussfolgerungen zulässt. Besitzt die Steuerbarkeitsmatrix eines Systems 2. Ordnung einen doppelten Singulärwert, so ist der größte Abstand zwischen zwei benachbarten Singulärwerten gleich Null. Dadurch würde fälschlicherweise angezeigt, dass das System nicht steuerbar ist.

Da in vielen Fällen gilt $\frac{\|\mathbf{A}_B\|}{\sigma_{r-1}} \gg 1$, ist für eine Abschätzung neben der Norm von \mathbf{A}_B auch das kleinste Verhältnis von zwei benachbarten Singulärwerten entscheidend.

3.2.2.5 Bestimmung einer oberen Schranke mittels Stufenform

Durch eine orthogonale Transformation kann ein beliebiges System (\mathbf{A}, \mathbf{B}) in die sogenannte Stufenform

$$\left[\mathbf{U}^H \mathbf{B} \mid \mathbf{U}^H \mathbf{A} \mathbf{U} \right] = \left[\begin{array}{c|cccc|c} \mathbf{X}_1 & \mathbf{A}_{1,1} & \mathbf{A}_{1,2} & \cdots & \mathbf{A}_{1,k} & \mathbf{A}_{1,k+1} \\ \mathbf{0} & \mathbf{X}_2 & \mathbf{A}_{2,2} & & \vdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & & \ddots & \mathbf{X}_k & \mathbf{A}_{k,k} & \mathbf{A}_{k,k+1} \\ \hline 0 & \cdots & \cdots & 0 & \mathbf{X}_{k+1} & \mathbf{A}_{k+1,k+1} \end{array} \right]$$

mit

$$\mathbf{A}_{i,i} \in \mathbb{C}^{p_i \times p_i} \text{ und } \mathbf{X}_i \in \mathbb{C}^{p_i \times p_{i-1}} \quad \text{für } 1 \leq i \leq k$$

überführt werden. Die Matrizen $\mathbf{A}_{i,i}$ und \mathbf{X}_i haben vollen Zeilenrang p_i und es gilt

$$\mathbf{A}_{k+1,k+1} \in \mathbb{C}^{(n-\sigma_k) \times (n-\sigma_k)} \quad \text{mit} \quad \sigma_k = \sum_{i=1}^k p_i.$$

Für Informationen zur Bestimmung der Transformationsmatrix \mathbf{U} siehe z.B. [7]. Da es sich hierbei um eine orthogonale Transformation handelt, gilt

$$\mu(\mathbf{A}, \mathbf{B}) = \mu(\mathbf{U}^H \mathbf{A} \mathbf{U}, \mathbf{U}^H \mathbf{B}).$$

Liegt das System in Stufenform vor, kann das Vorhanden- bzw. nicht Vorhandensein der Steuerbarkeit direkt abgelesen werden. Ist das System nicht steuerbar, dann ist $\mathbf{X}_{k+1} = \mathbf{0}$, und die Matrix $\mathbf{A}_{k+1,k+1}$ beschreibt den nicht steuerbaren Systemanteil. Das System ist genau dann steuerbar, wenn alle Matrizen \mathbf{X}_i den vollen Rang haben. Daraus ergibt sich eine einfache obere Schranke [48]

$$\mu_R \leq \min_i \|\mathbf{X}_i\|_2.$$

Für Eingrößensysteme entspricht die Stufenform der HESSENBERG-Form und die Matrizen \mathbf{X}_i sind Skalare x_i . Für die obere Schranke gilt in diesem Fall [48]

$$\mu_R \leq \min_i |x_i|.$$

Ein von MIMINIS und PAIGE in [36] angeführtes Beispiel zeigt, dass sich die Stufenform vor allem für große n zur Bestimmung des Distanzmaßes μ_R nicht eignet. D.h., es kann vorkommen, dass trotz hoher x_i Werte das „wahre“ μ_R sehr klein ist. Dies liegt vermutlich an der besonderen Struktur, die ein System in Stufenform aufweist. Große x_i Werte scheinen zwar darauf hinzuweisen, dass eine große Störung nötig ist, um eines dieser Elemente zu Null zu machen, jedoch zerstört eine sehr kleine Störung an der Stelle, an der sich eine Null befinden muss, die Struktur der Stufenform. Wird nun das gestörte System auf Stufenform gebracht, so kann man sich leicht vorstellen, dass durch die Transformation die x_i Werte groß werden. Je größer n , desto größer ist die Auswirkung dieses Effekts, da immer mehr Elemente unverändert bleiben müssen, um die Struktur der Matrizen zu erhalten. Aus diesem Grund liefert die durch die Stufenform ermittelte obere Schranke, wie später in den Beispielen gezeigt wird, oft Werte, die viel größer als μ_R sind.

3.2.3 Einfache Methoden zur Abschätzung von μ

Da die Berechnung des Distanzmaßes recht kompliziert ist, wird anhand von Beispielen die Aussagekraft der einfach zu bestimmenden oberen Schranke

$$\mu = \inf_{s \in \mathbb{C}} \{\sigma_{\min}([s\mathbf{E} - \mathbf{A}, \mathbf{B}])\} \leq \min\{\sigma_{\min}([\lambda_i \mathbf{E} - \mathbf{A}, \mathbf{B}])\} =: \mu_\lambda \quad (3.29)$$

mit λ_i , den Eigenwerten der Matrix \mathbf{A} , untersucht.

Im Allgemeinen können die Singulärwerte der Steuerbarkeitsmatrix \mathbf{S}_u nicht direkt zur Bestimmung des Distanzmaßes herangezogen werden (siehe Abschnitt 3.2.2.4). Da die Singulärwerte jedoch numerisch gut bestimmt werden können, werden der kleinste Singulärwert der Steuerbarkeitsmatrix $\sigma_{\min}(\mathbf{S}_u)$ und auch das kleinste Verhältnis von zwei benachbarten Singulärwerten bestimmt und deren Aussagekraft untersucht.

3.2.4 Überblick über die Methoden zur Ermittlung von μ

Abbildung 3.11 zeigt eine Zusammenfassung der vorgestellten Methoden zur Bestimmung des Distanzmaßes μ .

Die auf den Optimierungsproblemen basierenden Ansätze liefern obere und untere Schranken, deren Differenz vorgebar ist. Wird allerdings der gewünschte maximale Abstand zwischen oberer und unterer Schranke zu klein gewählt, so kann es auf Grund von numerischen Fehlern zu falschen Ergebnissen führen. Die Ansätze gehen von der Beziehung

$$\mu \leq \delta := \sigma_{\min}([s\mathbf{E} - \mathbf{A}, \mathbf{B}]) \quad (3.30)$$

aus. Somit ist δ der kleinste Singulärwert der Matrix $[s\mathbf{E} - \mathbf{A}, \mathbf{B}]$, weshalb die Singulärwertgleichungen (3.18) erfüllt sein müssen. Die Gleichungen führen zu einem Gleichungssystem bestehend aus drei vektoriellen Gleichungen. Besitzt dieses Gleichungssystem eine Lösung, dann ist Gleichung (3.30) erfüllt. GAO und NEUMANN formen das Gleichungssystem durch Variablensubstitution in ein Eigenwertproblem niedrigerer Ordnung um. Der Nachteil dieser Methode ist, dass nicht nur ein Wert für δ , sondern auch ein Wert für s vorgegeben werden muss, um das Gleichungssystem bzw. Eigenwertproblem auswerten zu können. Deshalb ist eine Rasterung des zu durchsuchenden Gebietes notwendig. Auch GU formt das Gleichungssystem in Eigenwertprobleme um. Die von GU vorgeschlagene Umformung ist deutlich aufwendiger als jene von GAO und NEUMANN. GU nutzt Theorem 5, welches ermöglicht, den Wert s zu berechnen, um die Gültigkeit von (3.30) zu überprüfen. Benutzt man für GUs Test die Umformung nach GAO und NEUMANN, kann die Berechnungsdauer deutlich verkürzt werden (siehe Abschnitt 3.2.6). Sowohl für die Methode nach GAO und NEUMANN als auch für die Methode nach GU existieren reelle Varianten, um das globale Minimum auf der reellen Achse und somit das Distanzmaß μ_R zu bestimmen.

Eine Ausnahme von den auf den Optimierungsproblemen basierenden Ansätzen ist der Algorithmus nach WICKS und DeCARLO, welcher keine Schranken, sondern einen Schätzwert für μ_C liefert. Im Gegensatz zu den übrigen Methoden arbeitet dieser Algorithmus mit der Idee der parallelen Lösung von zwei Optimierungsproblemen. Diese Methode kann nicht so einfach für reelle Störungen bzw. für die Berechnung von μ_R herangezogen werden.

Die übrigen vorgestellten Methoden nutzen unterschiedlichste Ansätze, um obere und untere Schranken für das Distanzmaß zu bestimmen. Der Abstand zwischen oberer und unterer Schranke kann hier nicht gewählt werden.

3.2.5 Überprüfung der Forderungen

Im Folgenden werden die Distanzmaße anhand der bereits bei den modalen Maßen in Abschnitt 2.4 verwendeten einfachen Beispiele auf Invarianz gegenüber Koordinatentransformationen, Konsistenz und Stetigkeit überprüft.

Beispiel 3.2 (Fortsetzung von Beispiel 2.2): Die Systemmatrizen seien gegeben durch

$$\mathbf{A} = \begin{bmatrix} \alpha & 1 \\ 0 & -2 \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

mit einem doppelten Eigenwert bei $\alpha = -2$ und einem Verlust der Steuerbarkeit für den Wert $\alpha = -3$.

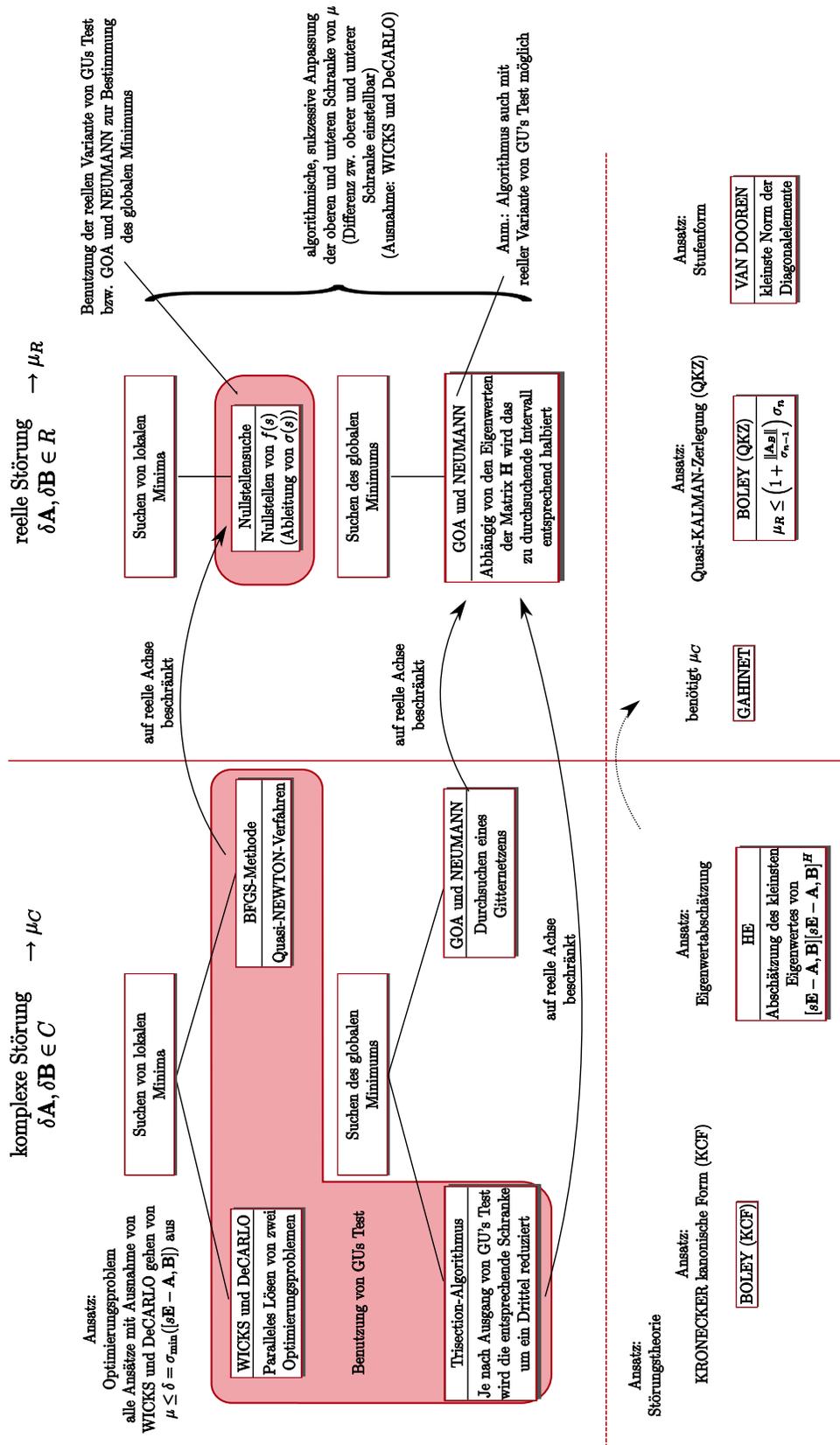


Abbildung 3.11: Übersicht über die numerische Berechnung von Schranken für das Distanzmaß μ

Die durch die vorgestellten Algorithmen bestimmten Schranken von μ_C haben i.A. ähnliche Verläufe und geben den Steuerbarkeitsverlust korrekt wieder. Der Algorithmus nach GAO und NEUMANN liefert eine weder stetige noch konsistente Schätzung für μ_C , da ein Gitternetz benutzt wird. Das Minimum kann somit nur dann exakt bestimmt werden, wenn es genau auf einem Gitterpunkt liegt.

Sehr enge Schranken (mit $tol = 10^{-9}$) liefert der Trisection-Algorithmus, dessen Ergebnisse in Abbildung 3.12 dargestellt sind. Ähnlich enge Schranken liefert auch die BFGS-Methode. Abbildung 3.12 zeigt die mittels der KRONECKER-Form erhaltenen Schranken, die deutlich weiter auseinander liegen. Das den Ergebnissen überlagerte „Rauschen“ kann durch die Verkleinerung der Elemente in der Matrix C in Gleichung (3.27) verbessert, jedoch nicht gänzlich beseitigt werden.

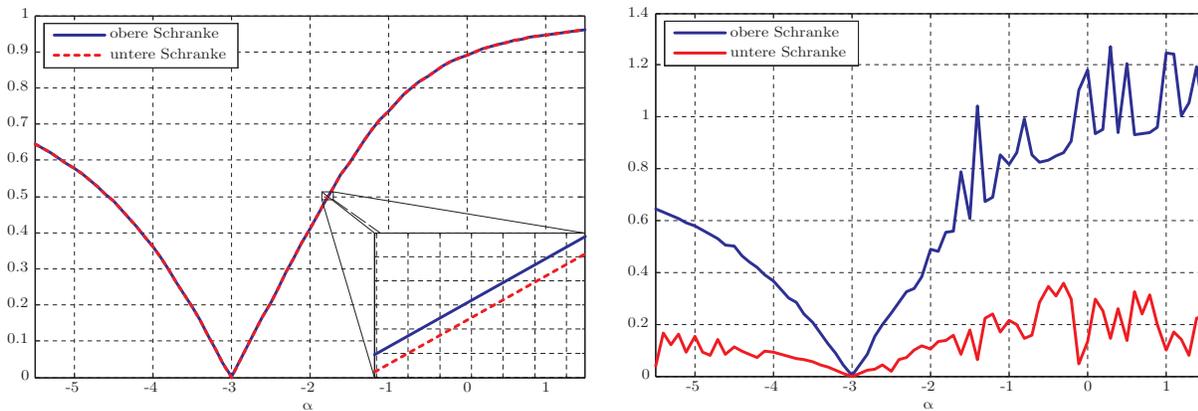


Abbildung 3.12: Mittels Trisection-Algorithmus (links) und aus KRONECKER-Form (rechts) erhaltene Schranken für μ_C

Der auf Eigenwertabschätzung basierende Ansatz nach HE aus Abschnitt 3.2.1.7 liefert bei dem doppelten Eigenwert $\alpha = \lambda_{1,2} = -2$ kein eindeutiges Ergebnis. Obwohl das System an dieser Stelle steuerbar ist, wird die untere Schranke zu Null, wodurch nicht mehr mit Sicherheit bestimmt werden kann, dass das System steuerbar ist.

Alle Ansätze zur Bestimmung des Distanzmaßes μ_R geben den Verlust der Steuerbarkeit korrekt wieder. Aus dem Pseudospektrum ist ersichtlich, dass das gesuchte Minimum auf der reellen Achse liegt, daher gilt $\mu_C = \mu_R$. Dieser Sachverhalt wird nur von dem Algorithmus nach GAO und NEUMANN (beschrieben in Abschnitt 3.2.2.2) sowie durch die Nullstellensuche von $f(s)$ (siehe Abschnitt 3.2.2.1) wiedergegeben. Die anderen Schätzungen für μ_R liefern zwar ähnliche Verläufe, sind aber ca. um Faktor 3 höher als μ_C .

In Abbildung 3.13 sind die Ergebnisse der in Abschnitt 3.2.3 vorgestellten einfachen Methoden zur Abschätzung von μ dargestellt. Zum Vergleich zeigt diese Abbildung zusätzlich das mittels Trisection-Algorithmus ermittelte μ_C . Der Wert μ_λ befindet sich in der gleichen Größenordnung wie μ_C und ist der einzige Wert, der eine obere Schranke darstellt. Auch die übrigen Werte befinden sich in der Größenordnung von μ_C . Alle Abschätzungen zeigen den Steuerbarkeitsverlust für $\alpha = -3$ korrekt an.

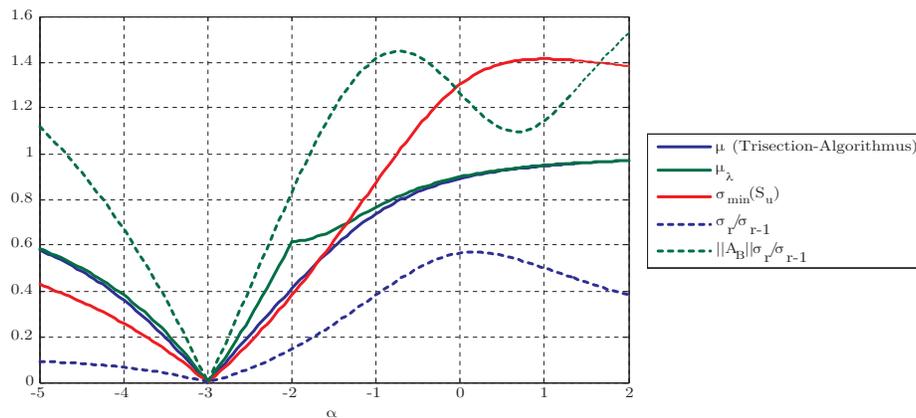


Abbildung 3.13: Einfache Methoden zur Bewertung der Steuerbarkeit

Beispiel 3.3 (Fortsetzung von Beispiel 2.3): Gegeben sei das in Modalform vorliegende System

$$\mathbf{A} = \begin{bmatrix} -1 & 0 \\ 0 & -1 + \alpha \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Im Gegensatz zu den Modalmaßen liefern die Distanzmaße konsistente Ergebnisse wegen des durch innere Verkopplung hervorgerufenen Steuerbarkeitsverlustes auf Grund des doppelten Eigenwerts $\lambda_{1,2} = -1$ an der Stelle $\alpha = 0$. Auch für dieses Beispiel gilt $\mu_C = \mu_R$. Die Algorithmen zur Bestimmung von μ_C und μ_R sind, mit Ausnahme der Berechnung von μ_C nach GAO und NEUMANN, konsistent und liefern ähnliche gute Ergebnisse wie im vorigen Beispiel. Im Falle der Bestimmung von μ_R sind bei diesem Beispiel nur die durch die Quasi-KALMAN-Zerlegung und die durch die Formel nach GAHINET erhaltenen Werte um Faktor 2 zu groß. Auch für dieses Beispiel erhält man für die einfachen Abschätzungen die gleichen Erkenntnisse wie schon beim vorherigen Beispiel:

Es gilt:

- Alle Abschätzungen liegen in der Größenordnung von μ_C und zeigen den Steuerbarkeitsverlust an.
- Es gilt $\mu_C \leq \mu_\lambda$.

Die Eigenschaften der Distanzmaße sind in den Tabellen 3.1 und 3.2 zusammengefasst. Da die Distanz zu dem nächsten nicht steuerbaren System selbst nur invariant gegenüber unitären Transformationen ist, wird bei den Maßen natürlich nur diese Eigenschaft überprüft. Im Falle von reellwertigen Störungen ist die Distanz μ_R , wie schon in Abschnitt 3.2.2 erläutert, nicht stetig, weshalb auch die Maße unstetig sein müssen.

3.2.6 Berechnungsdauer der Algorithmen

Da bei der Bestimmung von μ_C und μ_R auch die Berechnungsdauer der Algorithmen von Interesse ist, wurde diese mit dem Matlab Profiler gemessen und in Tabelle 3.3 zusammengestellt. Es wird ein mit Zufallszahlen gefülltes Eingrößensystem der Ordnung $n = 10$ herangezogen.

Tabelle 3.1: Vergleich der Distanzmaße zur Bestimmung von μ_C

	Invariant gegenüber unitären		
	Koordinatentrans.	Konsistenz	Stetigkeit
Trisection mit GUs Test	✓	✓	✓
BFGS mit GUs Test	✓	✓	✓
KRONECKER-Form	✓	✓	✓
Eigenwertabschätzung nach HE	x	✓	✓
WICKS und DeCARLO	x	✓	x
GAO und NEUMANN	✓	x	x

Tabelle 3.2: Vergleich der Distanzmaße zur Bestimmung von μ_R

	Invariant gegenüber unitären	
	Koordinatentrans.	Konsistenz
Formel nach GAHINET	✓	✓
Nullstellen von $f(s)$	✓	✓
Quasi-KALMAN-Zerlegung	✓	✓
Stufenform	✓	✓
GAO und NEUMANN	✓	✓

Bei der Berechnungsdauer der Algorithmen ist nicht die absolute Zeit von Interesse, da diese natürlich stark vom verwendeten Rechner abhängig ist. Die ermittelte Berechnungsdauer dient als Vergleich der Algorithmen untereinander.

Tabelle 3.3: Vergleich der Berechnungsdauer (für 50 Durchläufe)

		absolute Zeit [s]	Zeit [%] *	Bemerkung
μ_C	Trisection mit GUs Test	360,908	92,0	99% GUs Test
	BFGS mit GUs Test	25,681	6,5	98,2% GUs Test, 1,6% BFGS
	GAO und NEUMANN	5,475	1,4	
	WICKS und DeCARLO	0,219	< 1	
	Eigenwertabschätzung nach HE	0,159	< 1	
	KRONECKER-Form	0,020	< 1	
μ_R	Nullstellensuche von $f(s)$	0,930	60,0	97% Bisection, 3% Überprüfung nach GAO und NEUMANN
	GAO und NEUMANN	0,254	16,4	
	Quasi-KALMAN-Zerlegung	0,225	14,5	
	Stufenform	0,141	9,1	

* Bezogen auf die benötigte Berechnungsdauer aller Algorithmen für die Berechnung von μ_C bzw. μ_R .

Tabelle 3.3 zeigt, dass der Trisection-Algorithmus deutlich mehr Zeit benötigt als die übrigen Methoden. Hierbei beansprucht GUs Test aus zwei Gründen einen Großteil der Zeit (99%). Einerseits benötigt der Test an sich viel Zeit, da bei jedem Aufruf ein Eigenwertproblem der Größe $2n^2 \times 2n^2$ gelöst sowie eine QR-Zerlegung durchgeführt werden. Andererseits wird der Test bei jeder Reduktion der oberen oder unteren Schranke aufgerufen. In dem hier für 50 Systeme durchgeführten Beispiel wurde GUs Test 458 Mal zur Berechnung des Distanzmaßes benötigt. Auch bei der BFGS-Methode benötigt GUs Test einen großen Anteil der Gesamtzeit (98%). In diesem Fall wird er jedoch nur für die Entscheidung benötigt, ob ein globales oder lokales Minimum gefunden wurde. Somit wurde GUs Test nur 88 Mal ausgeführt, weshalb diese Methode deutlich schneller ist als der Trisection-Algorithmus.

Bei der Methode nach GAO und NEUMANN muss für die Entscheidung wie das Gitternetz durchlaufen wird, ein Eigenwertproblem gelöst werden. Dieses wird zwar 3964 Mal durchgeführt, jedoch hat das Problem nur eine Größe von $2n \times 2n$.

Die Berechnungsdauer der übrigen Ansätze ist sehr gering.

Die Berechnungsdauer der Algorithmen zur Bestimmung von μ_R ist deutlich kleiner als jene bei der Bestimmung von μ_C .

Von diesen benötigen die Nullstellensuchen von $f(s)$ deutlich mehr Zeit als die übrigen Ansätze. Den Hauptteil der Zeit benötigt der Algorithmus zur Suche der Nullstellen (97%) bzw. zur Auswertung der Funktion $f(s)$ an den Stützstellen. Die Funktion $f(s)$ wird 4158 Mal ausgewertet, wofür jeweils eine Singulärwertzerlegung notwendig ist.

3.2.6.1 Verbesserungen

Bei der Bestimmung des Distanzmaßes μ_C mit Hilfe des Trisection-Algorithmus bzw. der BFGS-Methode benötigt GUs Test einen Großteil der Zeit. Die Umformung in GUs Test kann durch jene nach GAO und NEUMANN ersetzt werden (siehe [19]). Damit ist keine QR-Zerlegung notwendig, und das generalisierte Eigenwertproblem wird zu einem einfachen Eigenwertproblem. Dies führt zu einer erheblichen Zeitersparnis bei der Berechnung des Distanzmaßes, wie nachfolgende Tabelle zeigt.⁶

Tabelle 3.4: Umformung nach GU vs. Umformung nach GAO und NEUMANN (für 50 Durchläufe)

Algorithmus	Umformung nach	
	GU	GAO und NEUMANN
Trisection	74%	26%
BFGS	68%	32%

Für eine weitere Beschleunigung der Algorithmen wird ausgenutzt, dass die reelle Variante von GUs Test deutlich schneller ist als die komplexe. Ist für einen Wert δ zu prüfen, ob die Ungleichung $\delta \geq \mu_C$ erfüllt ist, wird zuerst mit der reellen Variante von GUs Test überprüft, ob $\delta \geq \mu_R$ gilt. Ist das der Fall, gilt wegen $\mu_R \geq \mu_C$ auch $\delta \geq \mu_C$. Andernfalls muss die Ungleichung mit der komplexen Variante von GUs Test überprüft werden. Die Ergebnisse sind in nachfolgender Tabelle dargestellt.⁶Für die Berechnungen wurde GUs Test mit der Umformung nach GAO und NEUMANN benutzt.

Für den Trisection-Algorithmus ergibt sich durch diese Änderung eine erhebliche Verkürzung der Berechnungsdauer. Die komplexe Variante von GUs Test wurde 453 Mal aufgerufen, wäh-

Tabelle 3.5: Benutzung der reellen Variante von GUs Test (für 50 Durchläufe)

Algorithmus	ohne	mit
	Ausnutzung der reellen Variante	
Trisection	61%	39%
BFGS	50,8%	49,2%

rend diese bei der ursprünglichen Version des Algorithmus 738 Mal benutzt wurde. Bei der BFGS-Methode ist die Verbesserung sehr gering, da GUs Test im Vergleich zum Trisection-Algorithmus weniger oft zum Einsatz kommt. Die komplexe Variante wurde 87 anstatt 88 Mal verwendet.

Nachfolgend ein erneuter Vergleich der Berechnungsdauer der Algorithmen zur Ermittlung von μ_C nun mit den beiden Verbesserungen.

Tabelle 3.6: Vergleich der Berechnungsdauer für μ_C (für 50 Durchläufe)

	absolute Zeit [s]	Zeit [%]	Bemerkung
Trisection mit GUs Test	99,599	85,0	99% GUs Test
BFGS mit GUs Test	11,188	9,6	96,6% GUs Test, 2,8% BFGS
GAO und NEUMANN	5,475	4,7	
WICKS und DeCARLO	0,219	< 1	
Eigenwertabschätzung nach HE	0,159	< 1	
KRONECKER Form	0,020	< 1	

Tabelle 3.6 zeigt, dass der Trisection-Algorithmus und die BFGS-Methode immer noch mehr Zeit benötigen als die übrigen Methoden. Im Vergleich zu den ursprünglichen Versionen (siehe Tabelle 3.3) benötigt die verbesserte Version des Trisection-Algorithmus weniger als ein Drittel und die der BFGS-Methode weniger als die Hälfte der Zeit.

Durch die Umformung nach GAO und NEUMANN in GUs Test kann dieser deutlich beschleunigt werden. Jedoch wird immer noch relativ viel Zeit für das Lösen des Eigenwertproblems der Größe $2n^2 \times 2n^2$ benötigt. Motiviert durch die Tatsache, dass nur reelle Eigenwerte bestimmt werden müssen, werden in [19] zwei Verfahren vorgestellt, die die reelle Achse absuchen, um die gewünschten Eigenwerte zu finden. Dadurch kann die Anzahl der benötigten Operationen und damit die Berechnungsdauer von GUs Test weiter reduziert werden.

3.2.7 Diskussion der Ergebnisse

Auf Grund der beiden einfachen Beispiele können folgende Aussagen gemacht werden: Die auf dem Optimierungsproblem basierenden Ansätze liefern i.A. stetige und konsistente Ergebnisse. Nur auf die Methode nach GAO und NEUMANN zur Bestimmung des Distanzmaßes μ_C trifft das wegen der benötigten Rasterung des zu durchsuchenden Gebietes nicht zu.

Für die Bestimmung des Distanzmaßes μ_C liefern der Trisection-Algorithmus und die BFGS-

⁶Die Prozentangabe bezieht sich auf die Summe der Berechnungsdauer der beiden Varianten für den jeweiligen Algorithmus.

Methode identische Ergebnisse. Allerdings ist die Wahl des Toleranzwertes bei der BFGS-Methode kritischer. Wie erwähnt, kann der Test von GU ein falsches Ergebnis liefern, wenn der Abstand zwischen den Werten δ_1 und δ_2 zu klein ist. Dieser Abstand ist beim Trisection-Algorithmus sehr klein, wenn die obere und untere Schranke bereits eng beieinander liegen. Daher wirkt sich ein falsches Ergebnis von GUs Test nicht sehr stark auf das Ergebnis für das Distanzmaß μ_C aus. Wird jedoch bei der BFGS-Methode der Abstand zwischen δ_1 und δ_2 zu klein gewählt, so kann u.U. ein lokales Minimum als das globale Minimum interpretiert werden, was sich stark auf das Ergebnis für das Distanzmaß μ_C auswirken kann.

Diese beiden Methoden lassen sich sehr einfach auf die Minimumsuche auf der reellen Achse und somit zur Bestimmung des Distanzmaßes μ_R adaptieren. Für die hier durchgeführten Beispiele liefern die Nullstellensuchen von $f(s)$ und die reelle Variante der Methoden nach GAO und NEUMANN bzw. von GUs Test das identische konsistente und stetige Ergebnis.

Wie bereits in Abschnitt 3.2.2 erwähnt, kann der Ansatz des parallelen LöSENS von zwei Optimierungsproblemen von WICKS und DeCARLO nicht einfach zur Bestimmung von μ_R herangezogen werden. Die reelle Variante dieses Ansatzes führt nämlich zu einem deutlich schwierigeren Optimierungsproblem.

Der Ansatz der KRONECKER kanonischen Form und die Methode der Eigenwertabschätzung nach HE liefern im Vergleich zu den übrigen Methoden eine gute obere Schranke für μ_C . Beim Ansatz über die KCF ist die untere Schranke allerdings um einiges kleiner als das Distanzmaß μ_C . Hierbei ist das Ergebnis von der für die Berechnung benötigten mit Zufallszahlen gefüllten Matrix abhängig. Da der Ansatz der Eigenwertabschätzung nach HE auf der Modalform basiert, ist diese Abschätzung nicht anwendbar, wenn die Matrix \mathbf{A} mehrfache Eigenwerte besitzt. Die untere Schranke wird in diesem Fall zu Null, weshalb keine Aussage über die Steuerbarkeit des Systems möglich ist.

Der Ansatz über die Stufenform führt in manchen Fällen zu einer oberen Schranke, die deutlich höher ist als das Distanzmaß μ_C . Ein möglicher Grund dafür ist die besondere Struktur der Stufenform. Zur Bestimmung des Maßes werden nur ausgewählte Elemente dieser Struktur betrachtet. Die Norm der Störung, die diese Elemente zu Null macht (d.h. das System ist nicht mehr steuerbar), kann durchaus größer sein als die Norm der Störung, die auf andere Elemente der Matrix wirkt und dazu führt, dass das System nicht steuerbar wird.

Der Nachteil der Formel von GAHINET zur Bestimmung von μ_R besteht darin, dass vorab das Distanzmaß μ_C bestimmt werden muss.

Die Methoden für eine einfache Abschätzung liefern für die beiden Beispiele gute Ergebnisse, wobei nur der Wert μ_λ eine obere Schranke darstellt.

Kapitel 4

Anwendungen und Beispiele

In diesem Kapitel werden die Steuerbarkeitsmaße und deren Eigenschaften für ausgewählte Systeme weiteren mathematischen Untersuchungen unterzogen. In Abschnitt 4.1 werden die Maße für ein besonderes System in Steuerbarkeitsnormalform und für ein spezielles Mehrgrößensystem betrachtet. Weiters wird in Abschnitt 4.2 die Effizienz und Anwendbarkeit der unterschiedlichen Maßzahlen an praxisnahen Beispielen untersucht. In Abschnitt 4.3 werden mögliche Anwendungsgebiete der Maßzahlen erläutert. Hierbei wird auf die Wahl der Eigenwerte beim Zustandsreglerentwurf, auf die Fehlerdiagnose mit Hilfe von Steuerbarkeitsmaßen sowie auf die Bestimmung des optimalen Eingangsvektors bezüglich des Distanzmaßes näher eingegangen.

4.1 Beispiele

In Abschnitt 4.1.1 wird ein System in Steuerbarkeitsnormalform (SNF) betrachtet, für welches das Distanzmaß μ für beliebige Systemordnungen n analytisch angegeben werden kann. Einerseits werden dadurch wichtige Erkenntnisse über das Steuerbarkeitsverhalten dieser Systeme gewonnen. Andererseits können Aussagen über die Qualität der durch die verschiedenen Algorithmen und Ansätze ermittelten Schätzungen bzw. Schranken durch direkten Vergleich mit der analytischen Lösung gemacht werden. In Abschnitt 4.1.2 wird gezeigt, dass die Maße problemlos auch auf Mehrgrößensysteme anwendbar sind.

4.1.1 Steuerbarkeitsnormalform

Das in Steuerbarkeitsnormalform vorliegende System n -ter Ordnung mit den Systemdaten

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & & 0 \\ & 0 & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

entspricht der Serienschaltung von n Integratoren. KENNEY und LAUB haben in [24] die analytische Lösung des Distanzmaßes für diese Systeme

$$\mu_C = \sin\left(\frac{\pi}{n+1}\right)$$

bestimmt. Auffallend ist, dass mit anwachsender Systemordnung n das Distanzmaß kleiner, also die Steuerbarkeit des Systems schlechter wird. Die Singulärwerte der Steuerbarkeitsmatrix \mathbf{S}_u sind für beliebige n alle gleich Eins. Dadurch zeigt dieses Beispiel, dass die Singulärwerte von \mathbf{S}_u als Maß nicht geeignet sind, da sie den Verlauf des Steuerbarkeitsverhaltens i.A. nicht wiedergeben. Eine mögliche Erklärung für das Abnehmen der Steuerbarkeit mit zunehmender Systemordnung liefert die besondere Struktur dieser Systeme. Man kann sich leicht vorstellen, dass nur eine geringe Störung notwendig ist um die Struktur der Steuerbarkeitsnormalform zu verändern. Dadurch kann ein nicht steuerbares System entstehen. Dieser Effekt wird deutlicher, je größer die Systemordnung n ist, da mehr Elemente unverändert bleiben müssen, um die Struktur zu erhalten.

Aus Abbildung 4.1 ist ersichtlich, dass auch das Pseudospektrum eine außergewöhnliche Struktur aufweist. Die Funktion $\sigma_{\min}([s\mathbf{E} - \mathbf{A}, \mathbf{B}])$ nimmt an unendlich viele Stellen s den Wert des globalen Minimums an. Diese befinden sich auf einem Kreis mit dem Mittelpunkt $s = 0$, dessen Radius r von der Systemordnung n abhängt. Mit steigender Systemordnung wird der Radius größer und nähert sich dem Wert Eins an.

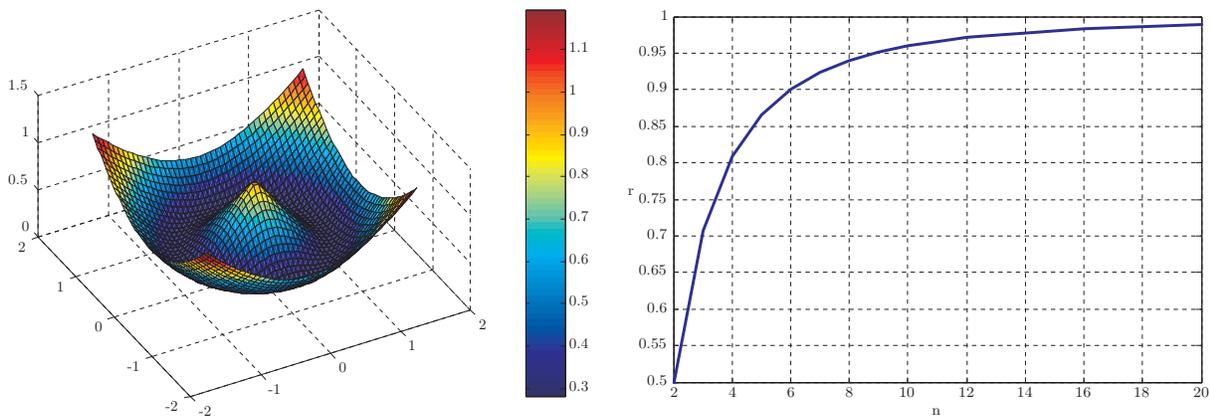


Abbildung 4.1: Pseudospektrum des Systems in SNF für $n = 10$ (links) und Abstand vom globalen Minimum zum Ursprung über Systemordnung n (rechts)

4.1.1.1 Distanzmaß μ_C

In Abbildung 4.2 wird die analytische Lösung für das Distanzmaß μ_C mit den aus den verschiedenen Methoden erhaltenen Schranken für μ_C verglichen.

Der Trisection-Algorithmus und die BFGS-Methode liefern sehr gute Schranken, die sich maximal durch den vorgegebenen Toleranzwert von der analytischen Lösung unterscheiden. Die Abweichung zwischen der analytischen Lösung und den durch den Algorithmus nach WICKS und DeCARLO ermittelten Schätzwertes ist kleiner als der vorgegebene Toleranzwert. Der Ansatz der KCF liefert eine gute obere Schranke. Die untere Schranke ist allerdings deutlich kleiner

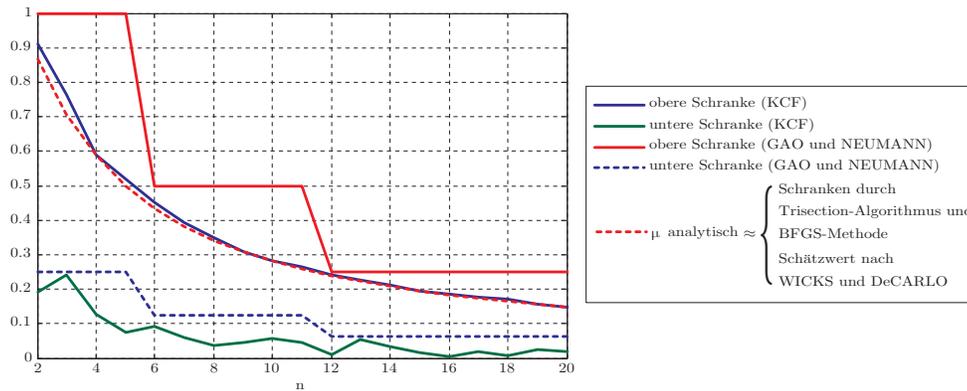


Abbildung 4.2: Distanzmaß μ des Systems in SNF über Systemordnung n^1

als die analytisch ermittelten Werte für μ_C .

Der Ansatz der Eigenwertabschätzung nach HE kann nicht angewendet werden, da mehrfache Eigenwerte vorliegen und diese Methode auf der Modalform basiert.

Auch an diesem Beispiel fällt bei der Methode nach GAO und NEUMANN auf, dass die berechneten Schranken stückweise konstant sind und sich meist deutlich von der analytischen Lösung unterscheiden. Dies liegt an der zu durchlaufenden Gitterstruktur. Die linke Seite der Abbildung 4.3 zeigt, den Ablauf des Algorithmus für $n = 1, \dots, 5$ (vgl. Abschnitt 3.2.1.5). Für diese Systemordnungen liegt der Wert des Distanzmaßes μ_C in dem mit A bezeichneten Rechteck. Das zu durchsuchende Gebiet kann gemäß Abschnitt 3 und auf Grund der Symmetrie auf $0 < \alpha \leq \|A\| = 1$ und $0 < s_o \leq \|A\| = 1$ beschränkt werden. Ausgangspunkt ist

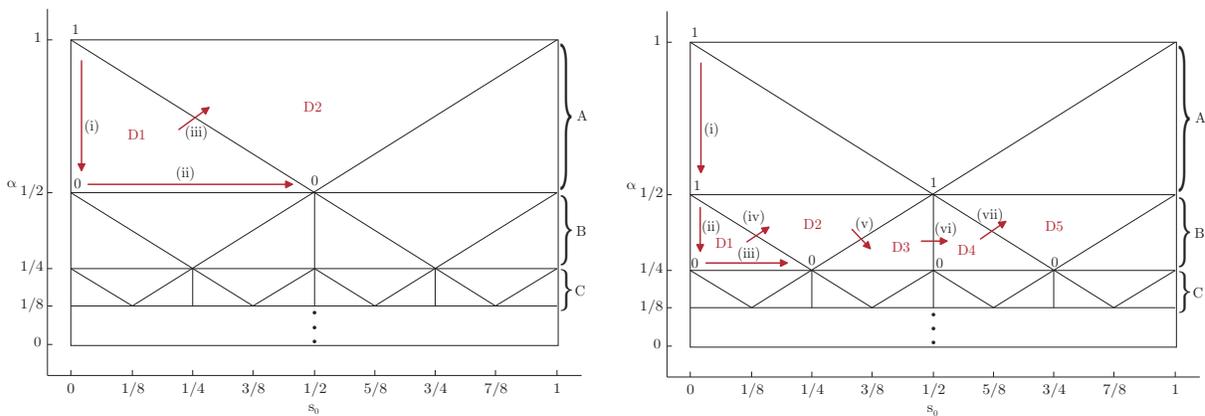


Abbildung 4.3: Algorithmus nach GAO und NEUMANN für $n = 1, \dots, 5$ (links) und für $n = 6, \dots, 11$ (rechts)

die linke obere Ecke $e_{1,1}$ mit den Koordinaten $(s_o = 0, \alpha = 1)$. Dieser Ecke wird der Wert $l(e_{1,1}) = 1$ zugewiesen, da $\alpha \geq \mu_C$ gilt. In Schritt (i) gelangt man durch Halbieren von α zur Ecke $e_{1,2} = (s_o = 0, \alpha = \frac{1}{2})$. Da $\alpha < \mu_C$ gilt, erhält diese Ecke den Wert $l(e_{1,2}) = 0$. In Schritt (ii) wird nun der Wert s_o verändert. Dadurch gelangt man zur Ecke $e_{2,2} = (\frac{1}{2}, \frac{1}{2})$ mit dem Wert

¹Für eine anschaulichere Darstellung wurden die ermittelten Werte durch Geraden verbunden.

$l(e_{2,2}) = 0$. Das Dreieck D1 besitzt eine DV-Kante (different value) $[e_{1,2} \ e_{2,2}]$, d.h. eine Kante deren Ecken verschiedene Werte zugeordnet sind. Diese Kante ist nun der Ausgangspunkt für das nächste zu betrachtende Dreieck D2 (Schritt (iii)). Die Ecke dieses Dreiecks, der noch kein Wert zugeordnet wurde ist der Punkt $e_{2,1}$. Da sich dieser Punkt am rechten Rand des zu durchsuchenden Dreiecks befindet ist die Suche hier zu Ende. Die obere Schranke wird nun durch das Minimum der α Werte jener Ecken, denen der Wert Eins zugeordnet wurde ermittelt. Für die Ordnungen $n = 1, \dots, 5$ gilt das nur für die Ecke $e_{1,1}$ mit dem Wert $\alpha = 1$.

Die rechte Seite der Abbildung 4.3 zeigt die Vorgangsweise des Algorithmus für die Ordnungen $n = 6, \dots, 11$. Hier liegt der Wert μ in Rechteck B. Ausgehend von der linken oberen Ecke $e_{1,1} = (0, 1)$ läuft die Suche durch Halbieren von α senkrecht nach unten bis zur Ecke $e_{1,3} = (0, \frac{1}{4})$ mit dem Wert $l(e_{1,3}) = 0$. Durch Verändern des Wertes s_0 gelangt man zur Ecke $e_{2,3} = (\frac{1}{4}, \frac{1}{4})$ und erhält die DV-Kante $[e_{2,3} \ e_{1,2}]$ die in Schritt (iv) ins Dreieck D2 führt. Wird nun der noch fehlenden Ecke dieses Dreiecks $e_{2,2}$ ein Wert zu gewiesen, erhält man wieder eine DV-Kante, die in das Dreieck D3 führt. Analog gelangt die Suche über weitere DV-Kanten in den Schritten (vi)-(vii) ins Dreieck D5. Die Ecke dieses Dreiecks liegt am Rand des Gebiets und die Suche wird beendet. Der kleinste α Wert der Ecken mit dem Wert Eins ist $\alpha = \frac{1}{2}$. Dies entspricht der durch den Algorithmus ermittelten oberen Schranken. An diesem Beispiel lassen sich einerseits die ermittelten konstanten Werte der Schranken als auch die große Abweichung der Schranken vom wahren Wert μ_C erklären. Ist der Wert μ_C groß, liegt dieser im oberen Bereich des zu durchsuchenden Gebiets, wo das Gitter noch sehr großmaschig ist und somit eine große Abweichung zustande kommt. Erst für kleine μ_C liefert der Algorithmus gute Schranken. Die Schranken können beispielsweise durch erneutes Anwenden des Algorithmus in einem kleineren Gebiet verbessert werden. Für die Ordnungen $n = 1, \dots, 5$ (linke Seite der Abbildung) würde das bedeuten, dass der Algorithmus auf das Rechteck A erneut angewendet werden muss. Analog kann der Algorithmus für die Ordnungen $n = 6, \dots, 11$ (rechte Seite der Abbildung) erneut auf das Rechteck B angewendet werden.

4.1.1.2 Distanzmaß μ_R

Zur Bestimmung des Distanzmaßes μ_R wird nur die reelle Achse betrachtet. Auf Grund der Struktur des Pseudospektrums ist ersichtlich, dass zwei reelle Werte s existieren, die das Optimierungsproblem minimieren. Somit gilt $\mu_R = \mu_C$. Dadurch kann die analytische Lösung auch zur Verifizierung der Ansätze zur Ermittlung von Schranken für das Distanzmaß μ_R genutzt werden. Der Algorithmus nach GAO und NEUMANN und die Nullstellen von $f(s)$ liefern Werte, die sich nur um den vorgegebenen Toleranzwert von der analytischen Lösung unterscheiden. Die übrigen Ansätze liefern weniger zufrieden stellende Ergebnisse. Die durch die Formel nach GAHINET ermittelte obere Schranke liefert zwar einen ähnlichen Verlauf wie die analytische Lösung, ist aber ca. um Faktor 2 größer. Der Ansatz über die Stufenform liefert den konstanten Wert Eins und die über die Quasi-KALMAN-Zerlegung ermittelte Schranke ist konstant gleich Zwei. Somit liefern diese beiden Ansätze gerade für Systeme hoher Ordnung Schranken, die deutlich größer sind als der wahre Wert des Distanzmaßes.

4.1.1.3 Einfache Methoden

Die in Abschnitt 3.2.3 vorgestellten einfachen Methoden liefern alle für beliebige Systemordnungen den Wert Eins. Einige dieser Methoden benutzen die Singulärwerte der Steuerbarkeits-

matrix bzw. das Verhältnis zwischen benachbarten Singulärwerten. Die Singulärwerte und somit die Verhältnisse sind für dieses Beispiel für beliebige Systemordnungen alle gleich Eins. Das Pseudospektrum ist an der Stelle $s = 0$ für beliebige Systemordnungen gleich Eins, weshalb auch die Schranke $\mu(\lambda) = \min\{\sigma(\lambda_i)\} = 1$ konstant ist.

4.1.2 Mehrgrößensystem

Im Folgenden wird gezeigt, dass die Maße ohne Schwierigkeiten auf Mehrgrößensysteme anwendbar sind. Hierfür dient das durch [35] motivierte akademische Beispiel

$$\mathbf{A} = \begin{bmatrix} 0,5 & 0,2 & 0,3 \\ 0,4 & 0,32 + \alpha & 0,3 \\ 0,2 & 0,5 & 0,26 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} 5 & 1 \\ 5 & 1 \\ 5 & 5 \end{bmatrix} \quad \mathbf{C} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

mit $n = 3$ Zustandsvariablen, $m = 2$ Eingangsgrößen und $p = 2$ Ausgangsgrößen. Für $\alpha = -0,02$ ist das System nicht steuerbar. Die Mindestanforderung an ein Maß ist, dass dieses den Steuerbarkeitsverlust anzeigt. Der Parameter α wird von $-0,05$ bis $0,02$ variiert und die Maße werden in Abhängigkeit dieses Parameters α dargestellt.

4.1.2.1 Modal- und Energiemaße

Alle Maße werden für mindestens einen Eigenwert bzw. für ein Eingangs-Ausgangspaar für $\alpha = -0,02$ gleich Null und geben somit den Steuerbarkeitsverlust korrekt wieder (vgl. Abbildung 4.4). Die Polempfindlichkeitsmaße S und s nach LITZ und das Regelbarkeitsmaß ρ nach HIPPE liefern $nmp = 12$ Werte, wodurch, wie in Abbildung 4.4 zu sehen, diese Maße für Mehrgrößensysteme sehr unübersichtlich werden.

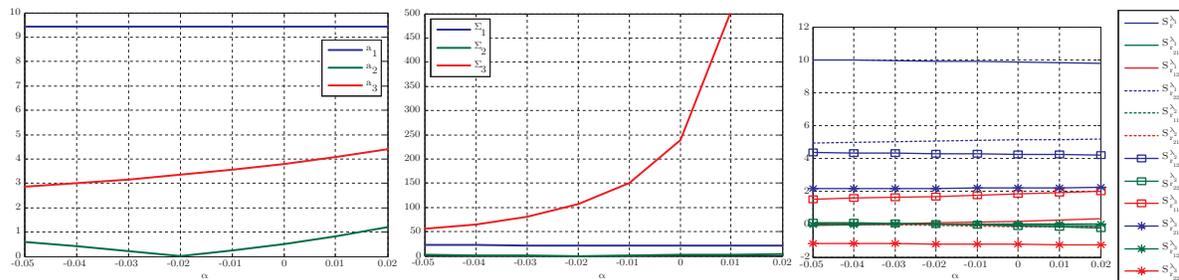


Abbildung 4.4: Modalmaße: Steuerbarkeitsmaß a nach KONNO (links), Summenmaß Σ (Mitte) und absolutes Polempfindlichkeitsmaß S (rechts) nach LITZ

Um die Maße bei Mehrgrößensystem für die verschiedenen Ein- und Ausgangsgrößen vergleichen zu können, sollten diese normiert werden. LITZ schlägt hierfür die Normierung mit dem betragsmäßig größten stationären Wert der Sprungantworten der Zustandsvariablen vor [29]. Da das hier untersuchte Beispiel positive Eigenwerte besitzt und damit instabil ist, ist diese Normierung in diesem Fall nicht möglich.

Auch die Energiemaße zeigen den Steuerbarkeitsverlust für $\alpha = -0,02$ an. Das Maß nach BENNINGER wird für die Zustandsvariablen x_1 und x_2 zu Null, was scheinbar im Widerspruch zu

den Modalmaßen steht, die anzeigen, dass ein Eigenwert nicht steuerbar ist. Betrachtet man die Transitionsmatrix des Systems für $\alpha = -0,02$

$$\Phi = \begin{bmatrix} 0,39e^t + 1,03e^{0,10t} - 0,42e^{-0,03t} & 0,33e^t - 1,03e^{0,10t} + 0,71e^{-0,03t} & 0,29e^t - 0,29e^{-0,03t} \\ 0,39e^t + 0,03e^{0,10t} - 0,42e^{-0,03t} & 0,33e^t - 0,03e^{0,10t} + 0,71e^{-0,03t} & 0,29e^t - 0,29e^{-0,03t} \\ 0,37e^t - 1,40e^{0,10t} + 1,03e^{-0,03t} & 0,31e^t + 1,40e^{0,10t} - 1,72e^{-0,03t} & 0,28e^t + 0,72e^{-0,03t} \end{bmatrix},$$

erkennt man, dass sich die Lösungen $x_1(t)$ und $x_2(t)$ für $x_1(0) = x_2(0)$ nicht voneinander unterscheiden. Da auch die Eingänge in gleicher Weise auf diese beiden Zustände wirken, können diese somit nicht beliebig beeinflusst werden und sind damit nicht steuerbar. Das Maß nach BENNINGER erlaubt somit einen besseren Einblick in das Steuerbarkeitsverhalten des Systems, da auch innere Verkopplungen berücksichtigt werden. Dies gilt nicht nur für mehrfache Eigenwerte (vgl. Beispiel 2.3) sondern wie dieses Beispiel zeigt, auch im Falle verschiedener Eigenwerte.

4.1.2.2 Distanzmaß μ_C

Alle in dieser Arbeit vorgestellten Algorithmen zur Bestimmung von μ_C bzw. von Schranken für μ_C zeigen den Steuerbarkeitsverlust für $\alpha = -0,02$ korrekt an. Der Trisection-Algorithmus und die BFGS-Methode liefern, wie in Abbildung 4.5 zu sehen, sehr enge Schranken. Alle übrigen Methoden liefern etwas weitere Schranken. Der Algorithmus nach WICKS und DeCARLO, der keine Schranken bestimmt, sondern μ_C direkt schätzt, liefert sehr gute Ergebnisse.

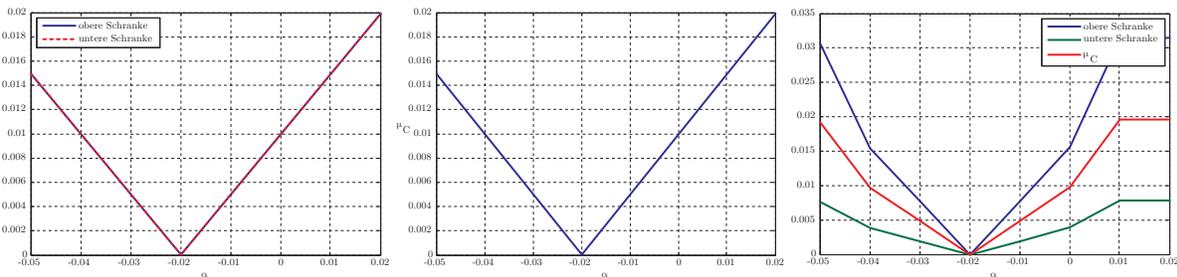


Abbildung 4.5: Distanzmaß μ_C : Trisection-Algorithmus und BFGS-Methode (links), Algorithmus nach WICKS und DeCARLO (Mitte) und Algorithmus nach GAO und NEUMANN (rechts)

4.1.2.3 Distanzmaß μ_R

Das Pseudospektrum sowie das mittels der BFGS-Methode ermittelte globale Minimum (vgl. Abbildung 4.6) zeigen, dass der Wert s^* , welcher Relation (3.3) minimiert, auf der reellen Achse liegt, d.h. $s^* \in \mathbb{R}$. Damit gilt $\mu_C = \mu_R$.

Abbildung 4.7 zeigt, dass die mittels des Algorithmus nach GAO und NEUMANN und die Nullstellensuche von $f(s)$ für μ_R gleich den Werten von μ_C sind. Für dieses Beispiel gilt das auch für die mittels Stufenform ermittelten Werte. Die durch die Eigenwertabschätzung nach HE und durch die Quasi-KALMAN-Zerlegung ermittelten oberen Schranken liefern etwas höhere Werte.

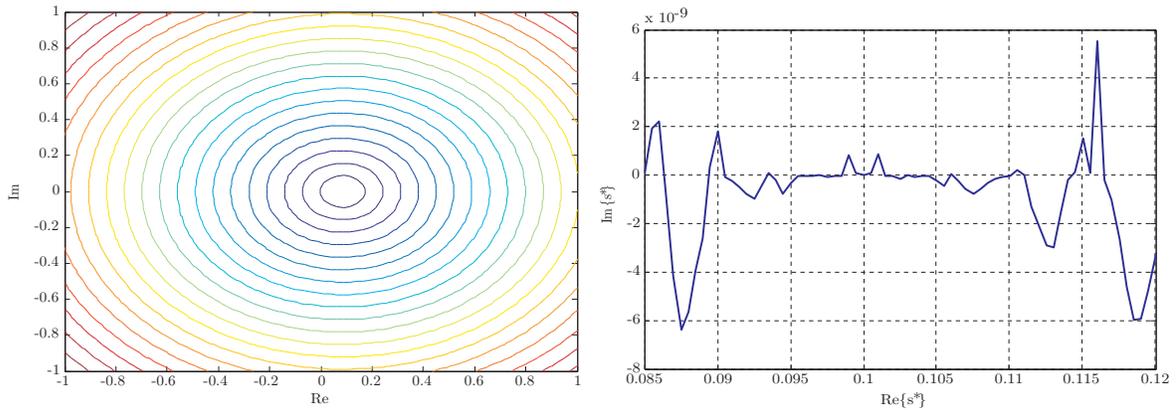


Abbildung 4.6: Pseudospektrum für $\alpha = -0,05$ (links), durch BFGS-Methode ermittelte Lage des globalen Minimums (rechts)

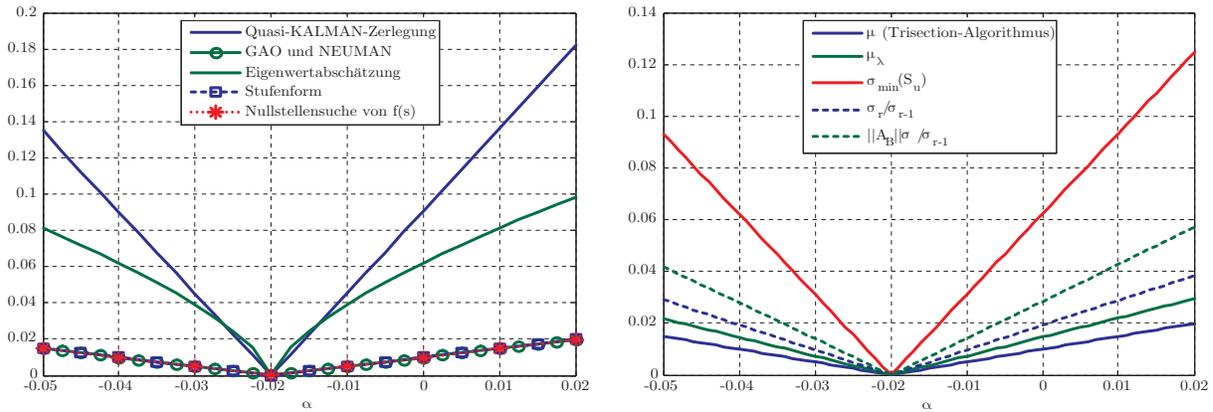


Abbildung 4.7: Distanzmaß μ_R (links) und einfache Methoden zur Berechnung von Schranken für μ_C (rechts)

4.1.2.4 Einfache Methoden

Wie aus Abbildung 4.7 ersichtlich, geben alle mittels der einfachen Methoden berechneten Schätzungen für das Distanzmaß den Steuerbarkeitsverlust korrekt an. Die Ergebnisse sind denen durch den Trisection-Algorithmus ähnlich, wobei der Wert μ_λ diesen Schranken am nächsten kommt.

4.2 Elektromechanische Beispiele

Nachfolgend werden für drei unterschiedliche Systeme Steuer- bzw. Beobachtbarkeitsmaße bestimmt. Durch die Anwendung der Maße auf mathematische Modelle von elektromechanischen Aufbauten können die Aussagen der Maße physikalisch interpretiert werden.

Als erstes Beispiel dient ein als „Wagen mit Stab“ bezeichneter Aufbau. Anhand dieses Aufbaus werden die notwendigen Schritte erläutert, die durchzuführen sind, um mit Hilfe der vorge-

stellten Maße Aussagen über die „Güte“ der Steuerbarkeit des betrachteten Systems zu treffen. Die erforderlichen Schritte lassen sich wie folgt unterteilen:

1. Modellbildung: Durch physikalische Zusammenhänge (NEWTONsche Axiome, KIRCHHOFFsche Gesetze, Bauelementgleichungen, Reibung, ...) erhält man i.A. nichtlineare Differentialgleichungen als mathematische Beschreibung des realen Systems.
2. Linearisierung: Da mit den hier vorgestellten Methoden nur die Steuerbarkeit von LZI-Systemen beurteilt werden kann, wird das nichtlineare System um den zu untersuchenden Arbeitspunkt linearisiert.
3. Berechnung der Maßzahlen.
4. Interpretation.

Als zweites Beispiel dient der sogenannte „Hubmagnet“. Hierbei handelt es sich um ein Labormodell, das am Institut für Regelungs- und Automatisierungstechnik an der Technischen Universität Graz aufgebaut ist.

Als drittes Beispiel wurde das „Doppelpendel“ gewählt. Dieser Aufbau wird durch ein LZI-System 6. Ordnung beschrieben. Es wird untersucht, ob die Effizienz und Interpretierbarkeit der Maße auch für Systeme höherer Ordnung gegeben ist. Als eine andere Herangehensweise bei der Modellbildung, wird anhand dieses Beispiels kurz erläutert, wie man mit Hilfe des LAGRANGE-Formalismus ein Modell erhält.

4.2.1 Wagen mit Stab

Im ersten Beispiel werden die Steuerbarkeitsmaße an einem sehr einfachen Modell, dem in Abbildung 4.8 dargestellten Wagen mit Stab, untersucht. Aus dem 2. NEWTONschen Axiom erhält man die Bewegungsgleichung des Wagens

$$\left(1 + \frac{m}{M}\right) \frac{d^2x}{dt^2} = \frac{F}{M} + \frac{m}{M} l \left(\frac{d\varphi}{dt}\right)^2 \sin \varphi - \frac{m}{M} l \frac{d^2\varphi}{dt^2} \cos \varphi \quad (4.1)$$

und mit dem Drallsatz die Differentialgleichung für die Bewegung des Stabes

$$\frac{d^2x}{dt^2} \cos \varphi + l \frac{d^2\varphi}{dt^2} = g \sin \varphi.$$

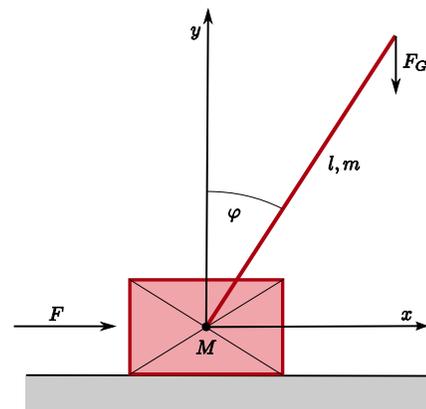


Abbildung 4.8: Wagen mit Stab

Unter der Annahme, dass die Masse des Pendels im Vergleich zum Wagen vernachlässigbar klein ist, d.h. $\left|\frac{m}{M}\right| \ll 1$, erhält man aus Gleichung (4.1)

$$\frac{d^2x}{dt^2} = \frac{F}{M}$$

Durch die Wahl der Zustandsvariablen

$$x_1 := \varphi, \quad x_2 := \frac{d\varphi}{dt}$$

und mit der Abkürzung

$$u := \frac{F}{M}$$

erhält man für die Bewegung des Stabes das Zustandsraummodell der Form $\frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x}, u)$

$$\begin{aligned} \frac{dx_1}{dt} &= x_2 \\ \frac{dx_2}{dt} &= \frac{g}{l} \sin x_1 - \frac{1}{l} u \cos x_1. \end{aligned}$$

Aus Gründen der Einfachheit gilt für die weiteren Berechnungen $l = 1$. Das System besitzt die Ruhelagen \mathbf{x}_R , für deren Komponenten gilt

$$x_{2,R} = 0 \quad \text{und} \quad u_R = g \tan x_{1,R}.$$

Für das linearisierte System

$$\frac{d\boldsymbol{\zeta}}{dt} = \mathbf{A}\boldsymbol{\zeta} + \mathbf{b}v$$

werden die JACOBI Matrizen in der Ruhelage

$$\mathbf{A} := \left. \frac{\partial \mathbf{f}}{\partial \mathbf{x}} \right|_{\mathbf{x}=\mathbf{x}_R, u=u_R} = \begin{bmatrix} 0 & 1 \\ g \cos x_{1,R} + u_R \sin x_{1,R} & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ \frac{g}{\cos x_{1,R}} & 0 \end{bmatrix}$$

$$\mathbf{b} := \left. \frac{\partial \mathbf{f}}{\partial u} \right|_{\mathbf{x}=\mathbf{x}_R, u=u_R} = \begin{bmatrix} 0 \\ -\cos x_{1,R} \end{bmatrix}$$

berechnet. Im Folgenden werden die Steuerbarkeitsmaße des linearisierten Modells im Intervall $x_{1,R} \in [-\pi, \pi]$ betrachtet. Die Matrix \mathbf{A} besitzt die Eigenwerte

$$\lambda_{1,2} = \pm \sqrt{\frac{g}{\cos x_{1,R}}}.$$

Ist $x_{1,R} \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$, nimmt der Kosinus positive Werte an, die Matrix \mathbf{A} besitzt zwei reelle Eigenwerte. Im restlichen Intervall ist der Kosinus negativ, und die Matrix besitzt somit ein rein imaginäres Eigenwertpaar. Für die Determinante der Steuerbarkeitsmatrix gilt

$$\det(\mathbf{S}_u) = \det([\mathbf{b}, \mathbf{A}\mathbf{b}]) = -\cos^2 x_{1,R}.$$

Für die Ruhelage $x_{1,R} = \pm \frac{\pi}{2}$, dies entspricht der waagrechten Lage des Stabes, wird die Determinante zu Null. Somit ist das System in diesen Punkten nicht steuerbar. Um den Stab in dieser Lage zu halten, muss die Eingangsgröße u_R unendlich groß sein. Obwohl diese Stablage damit in der Praxis nicht realisierbar ist, wird diese hier genauer untersucht um ein besseres Verständnis für die Maßzahlen und die Algorithmen zu deren Berechnung zu erhalten. In der Systemmatrix \mathbf{A} tritt für die waagrechte Stablage im Element $a_{2,1}$ eine Division durch Null auf. Deshalb gibt, wie anschließend gezeigt wird, nicht jedes Steuerbarkeitsmaß den Steuerbarkeitsverlust korrekt wieder.

4.2.1.1 Modal- und Energiemaße

Für das modaltransformierte System erhält man

$$\frac{d\zeta}{dt} = \sqrt{\frac{g}{\cos x_{1,R}}} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \zeta + \frac{(\cos x_{1,R})^{\frac{3}{2}}}{2\sqrt{g}} \begin{bmatrix} -1 \\ 1 \end{bmatrix} v.$$

Für die waagrechte Lage des Pendels wird der Eingangsvektor \mathbf{b} zum Nullvektor. Daraus erkennt man, dass die Modalmaße den Steuerbarkeitsverlust korrekt wiedergeben. Für diese Lage des Pendels wird ein Element der Systemmatrix \mathbf{A} unendlich groß. Auf Grund der Numerik ist jedoch $\cos(\pm\frac{\pi}{2})$ nicht exakt gleich Null. Dadurch ergibt sich trotzdem eine numerisch ermittelte Modalform, aus welcher sich für diesen Punkt Steuerbarkeitsmaße (mit Ausnahme von κ_s) kleiner 10^{-16} ergeben. In Abbildung 4.9 ist beispielhaft das Ergebnis für das Steuerbarkeitsmaß κ nach LITZ und a nach KONNO dargestellt. Die übrigen Maße, mit Ausnahme des Steuerbarkeitsmaßes κ_s nach LÜCKEL und MÜLLER, liefern einen ähnlichen Verlauf der Steuerbarkeit. Da die Realteile der Eigenwerte der Matrix \mathbf{A} für $x_{1,R} \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ betragsmäßig sehr hohe Werte annehmen, während sie im Rest des untersuchten Intervalls gleich Null sind, liefern die Steuerbarkeitsmaße κ_{si} auf Grund der Multiplikation mit $e^{-2\text{Re}\{\lambda\}}$ unbrauchbare Ergebnisse. Das Steuerbarkeitsmaß nach BENNINGER gibt an, dass die Zustandsvariable x_2 (Winkelgeschwindigkeit) besser steuerbar ist als x_1 (Winkel), was rein intuitiv nicht unmittelbar erkennbar ist.

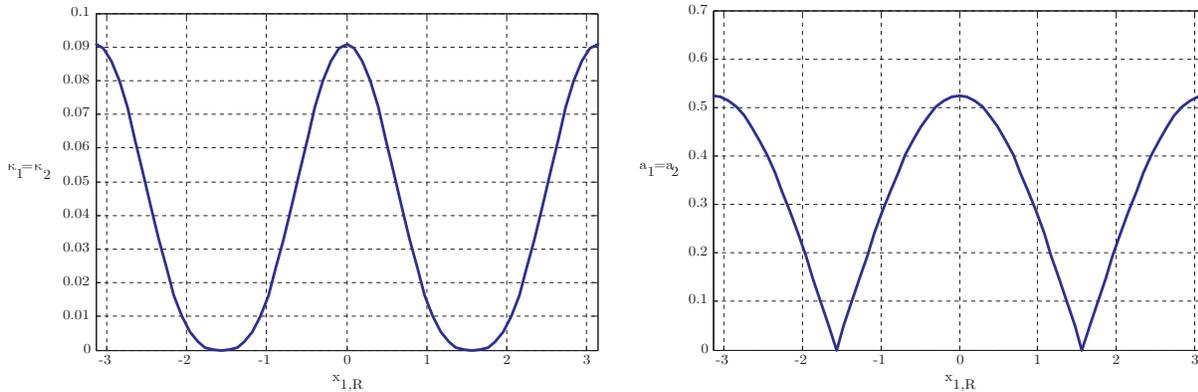


Abbildung 4.9: Steuerbarkeitsmaße κ und a für verschiedene Stabpositionen

4.2.1.2 Distanzmaß μ_C

Der Trisection-Algorithmus und die BFGS-Methode liefern gute und sehr eng beieinanderliegenden Schranken (mit $tol = 10^{-9}$), dargestellt in Abbildung 4.10. Auch der nach WICKS und DeCARLO ermittelte Schätzwert für μ_C liefert gute Ergebnisse. Alle auf dem Optimierungsproblem basierenden Algorithmen haben Schwierigkeiten an den Stellen $x_{1,R} = \pm\frac{\pi}{2}$. Dies entspricht der waagrechten Lage des Stabes für die das System nicht steuerbar ist. Die Maße streben zwar in der Umgebung dieser Unstetigkeitsstellen gegen Null, liefern aber für $x_{1,R} = \pm\frac{\pi}{2}$ Werte ungleich Null und somit falsche Ergebnisse. Dies kann folgendermaßen erklärt werden:

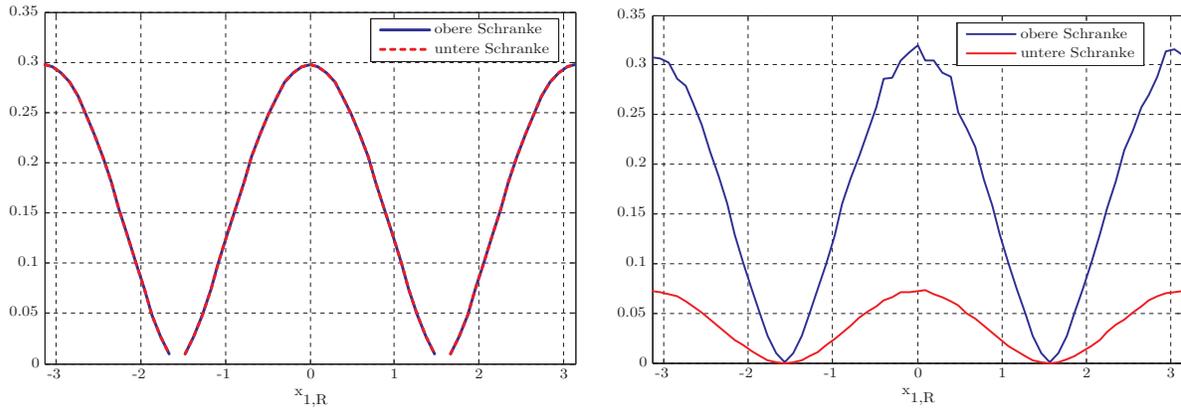


Abbildung 4.10: Bestimmung von μ_C mittels Trisection-Algorithmus (links) und der KRONECKER kanonischen Form (rechts) für verschieden Stabpositionen

Optimiert wird über den kleinsten Singulärwert der Matrix $\mathbf{M} := [s\mathbf{E} - \mathbf{A}, \mathbf{b}]$. Da die Singulärwerte der Matrix \mathbf{M} gleich den Quadratwurzeln der Eigenwerte von $\mathbf{M}\mathbf{M}^H$ sind, wird für die Berechnung die Beziehung

$$\det(z\mathbf{E} - \mathbf{M}\mathbf{M}^H) = 0$$

betrachtet. Ist das System nicht steuerbar, so ist der kleinste Singulärwert von \mathbf{M} gleich Null. Daraus folgt dass die Matrix $\mathbf{M}\mathbf{M}^H$ einen Eigenwert bei Null besitzen muss, d.h. $z = 0$ bzw.

$$\det(\mathbf{M}\mathbf{M}^H) = 0.$$

Es muss also ein Wert s existieren, der obige Bedingung für $x_{1,R} = \pm \frac{\pi}{2}$ erfüllt. Die Determinante lautet

$$\det(\mathbf{M}\mathbf{M}^H) = |s|^4 + |s|^2 \cos^2 x_{1,R} - \left(\frac{g}{\cos x_{1,R}} \right) (s^2 - \bar{s}^2) + \left(\frac{g}{\cos x_{1,R}} \right)^2 + \cos^2 x_{1,R}$$

Durch Einsetzen von

$$\tilde{s} = \sqrt{\frac{g}{\cos x_{1,R}}}$$

erhält man

$$\det(\mathbf{M}\mathbf{M}^H)|_{s=\tilde{s}} = \cos^2 x_{1,R} \left(\left| \frac{g}{\cos x_{1,R}} \right| + 1 \right).$$

Strebt der Winkel $x_{1,R}$ gegen $\pm \frac{\pi}{2}$ wird diese Determinante zu Null

$$\lim_{x_{1,R} \rightarrow \pm \frac{\pi}{2}} \det(\mathbf{M}\mathbf{M}^H)|_{s=\tilde{s}} = 0.$$

Der gesuchte Wert \tilde{s} strebt in diesem Fall gegen Unendlich. Das erklärt warum die numerische Suche nach diesem Wert nicht funktioniert und die darauf basierenden Algorithmen keine brauchbaren Ergebnisse liefern. Die durch die KRONECKER-Form und durch die Eigenwertabschätzung nach HE bestimmten Werte liefern gute obere Schranken. Jedoch sind die unteren Schranken deutlich kleiner als μ_C (siehe Abbildung 4.10). Mit beiden Methoden erhält man für

die untere Schranke bei waagrechtter Lage des Stabes den Wert Null. Der von GAO und NEUMANN vorgestellte Algorithmus gibt zwar den Verlauf der Steuerbarkeit i.A. richtig wieder, weist aber im betrachteten Intervall Sprungstellen auf. Ein Grund dafür ist das fest vorgegebene Gitter, welches über den zu durchsuchenden Teil des „Gebirges“ gelegt wird. Wie gut das Ergebnis ist, hängt davon ab, wie nahe sich das Minimum an einem Gitterpunkt befindet (vgl. Abschnitt 4.1.1.1).

4.2.1.3 Distanzmaß μ_R

Bei der Bestimmung von μ_R liefern der Algorithmus nach GAO und NEUMANN und die Nullstellensuche von $f(s)$ das gleiche Ergebnis (Abbildung 4.11).

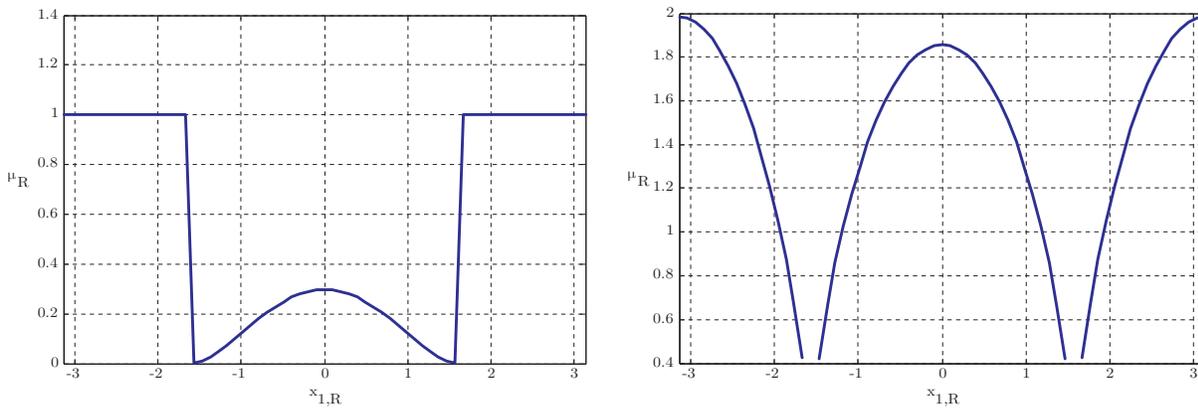


Abbildung 4.11: Bestimmung von μ_R mit dem Algorithmus nach GAO und NEUMANN sowie der Nullstellensuche von $f(s)$ (links) und durch die Formel nach GAHINET (rechts) für verschieden Stabpositionen

Der etwas eigentümliche Verlauf des durch die beiden Algorithmen bestimmten Maßes μ_R kann durch die Betrachtung des Pseudospektrums erklärt werden. Im Intervall $x_{1,R} \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ hat das Pseudospektrum zwei Minima auf der reellen Achse. Dadurch gilt in diesem Intervall $\mu_C = \mu_R$. Im Rest des Intervalls tritt ein konjugiert komplexes Minimum auf. In der Mitte davon befindet sich das Minimum auf der reellen Achse, welches konstant bleibt (vgl. Abbildung 4.12). Abbildung 4.11 zeigt auch die durch die Formel nach GAHINET bestimmten Werte μ_R , deren Verlauf denen von μ_C ähnelt, jedoch nicht das durch das Pseudospektrum zu sehende Verhalten wiedergibt.

Die Bestimmung von μ_R mittels der Quasi-KALMAN-Zerlegung liefert in der Umgebung von $x_{1,R} = \pm \frac{\pi}{2}$ sehr hohe Werte, wodurch fälschlicherweise gute Steuerbarkeit angezeigt wird. Das liegt daran, dass das Element $a_{21} = \frac{g}{\cos x_{1,R}}$ der Matrix \mathbf{A} für diese Werte von $x_{1,R}$ sehr groß wird, während das Element a_{12} konstant 1 bleibt. Dies führt dazu, dass die Norm der Begleitmatrix \mathbf{A}_B einen großen Wert annimmt und dass die Singulärwerte $\sigma_{1,2} = \cos x_{1,R}$ der Steuerbarkeitsmatrix \mathbf{S}_u in der Umgebung von $x_{1,R} = \pm \frac{\pi}{2}$ sehr klein werden, wodurch die großen Werte von μ_R in dieser Umgebung zustande kommen. Die Singulärwerte $\sigma_{1,2} = \cos x_{1,R}$ der Steuerbarkeitsmatrix sind Null, doch die numerische Auswertung von $\cos \pm \frac{\pi}{2}$ liefert einen sehr kleinen Wert, weshalb der Steuerbarkeitsverlust nicht korrekt wiedergegeben wird.

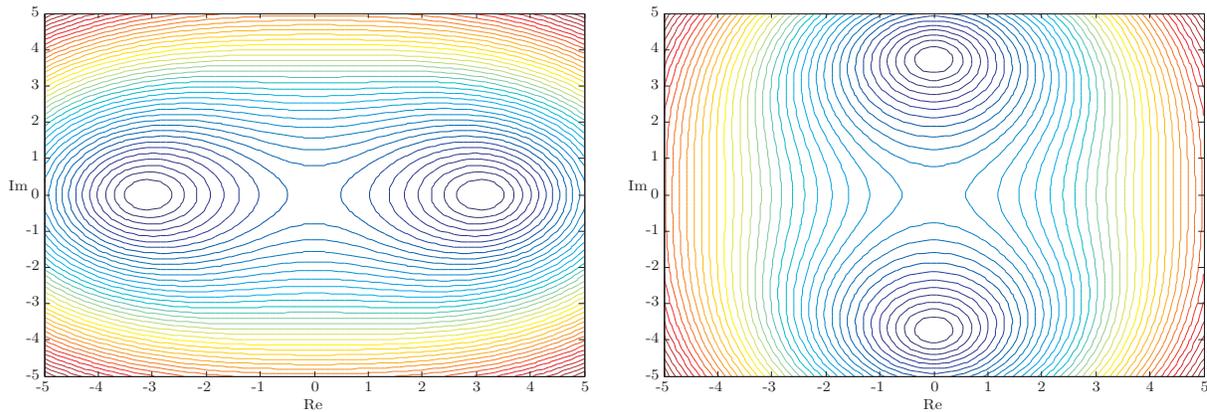


Abbildung 4.12: Pseudospektrum für $x_{1,R} = 0$ (links) und für $x_{2,R} = \frac{3\pi}{4}$ (rechts)

Das linearisierte System liegt in HESSENBERG Form vor, und das Element a_{21} ist immer größer gleich 1, weshalb die Bestimmung mittels Stufenform $\mu_R = 1$ für alle $x_{1,R}$ liefert.

4.2.1.4 Einfache Methoden

Für die obere Schranke μ_λ nach Gleichung (3.29) gilt, wie in Abbildung 4.13 zu sehen, $\mu_C \approx \mu_\lambda$. Der Verlauf des Singulärwerts der Steuerbarkeitsmatrix ähnelt dem von μ_C , liefert aber deutlich höhere Werte. Da die Steuerbarkeitsmatrix S_u einen doppelten Singulärwert besitzt, gilt $\frac{\sigma_r}{\sigma_{r-1}} = 1$ für alle Werte von $x_{1,R}$, womit diese Schranke in diesem Fall keine brauchbaren Ergebnisse liefert. Da die Norm $\|A_B\|$ für dieses Beispiel sehr groß ist, liefert auch die Abschätzung $\|A_B\| \frac{\sigma_r}{\sigma_{r-1}}$ Werte, die sehr viel größer sind als μ_C . Auch bei diesem Beispiel zeigen μ_λ und $\sigma_{\min}(S_u)$ den Steuerbarkeitsverlust richtig an.

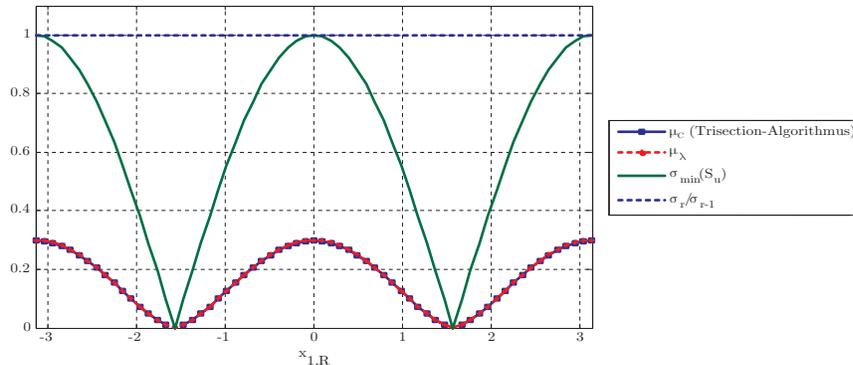


Abbildung 4.13: Einfache Methoden zur Abschätzung von μ

4.2.1.5 Betrachtung des nichtlinearen Systems

Ein Einblick in das Steuerbarkeitsverhalten von nichtlinearen Systemen kann durch eine mit Hilfe von LIE-Klammern erzeugte Steuerbarkeitsmatrix $\tilde{\mathbf{S}}_u$ erfolgen. Hierfür kann für ein sogenanntes AI-System (affine input)

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})u$$

die Steuerbarkeitsmatrix

$$\tilde{\mathbf{S}}_u = [\mathbf{g}, ad_{\mathbf{f}}\mathbf{g}, \dots, ad_{\mathbf{f}}^{m-1}\mathbf{g}]$$

herangezogen werden [42]. Hierbei bezeichnet $ad_{\mathbf{f}}\mathbf{g}$ die sogenannte *LIE-Klammer*, die folgendermaßen definiert ist:

Definition 2 (LIE-Klammer): Es seien \mathbf{f} und \mathbf{g} zwei Vektorfelder im \mathbb{R}^n . Die LIE-Klammer von \mathbf{f} und \mathbf{g} ist ein drittes Vektorfeld, beschrieben durch

$$ad_{\mathbf{f}}\mathbf{g} = \nabla\mathbf{g}\mathbf{f} - \nabla\mathbf{f}\mathbf{g}$$

mit

$$\nabla\mathbf{f}(\mathbf{x}) = \frac{\partial\mathbf{f}}{\partial\mathbf{x}},$$

der JACOBI-Matrix von \mathbf{f} .

Alternativ wird die LIE-Klammer auch als $[\mathbf{f}, \mathbf{g}]$ geschrieben. Das wiederholte Ausführen einer LIE-Klammer kann rekursiv durch

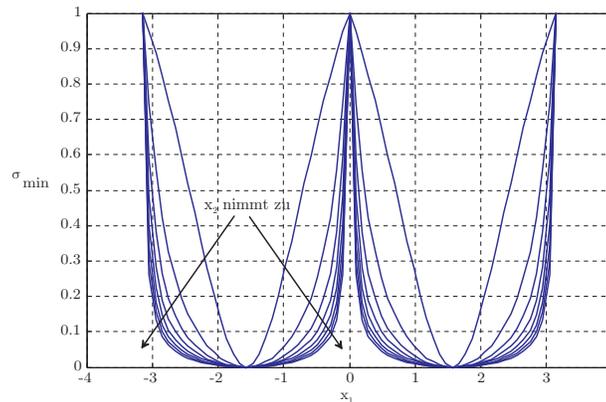
$$\begin{aligned} ad_{\mathbf{f}}^0\mathbf{g} &= \mathbf{g} \\ ad_{\mathbf{f}}^i\mathbf{g} &= [\mathbf{f}, ad_{\mathbf{f}}^{i-1}\mathbf{g}] \end{aligned}$$

definiert werden.

Berechnet man nun diese Steuerbarkeitsmatrix für das Beispiel „Wagen mit Stab“ so erhält man

$$\tilde{\mathbf{S}}_u = \begin{bmatrix} 0 & -\cos(x_1) \\ \cos(x_1) & -\sin(x_1)x_2 \end{bmatrix}.$$

Man sieht, dass im Gegensatz zum linearisierten System in diesem Fall die Steuerbarkeit auch von der Winkelgeschwindigkeit des Stabes $x_2 = \dot{\omega}$ abhängig ist. Als ein Hinweis auf gute oder schlechte Steuerbarkeit ist in Abbildung 4.14 der kleinste Singulärwert der Matrix $\tilde{\mathbf{S}}_u$ über $x_1 \in [-\pi, \pi]$ für unterschiedliche Werte von x_2 aufgetragen. Vergleicht man den Verlauf von $\sigma_{\min}(\mathbf{S}_u)$ mit dem Verlauf der Steuerbarkeitsmaße des linearisierten Systems (vgl. Abbildung 4.9), erkennt man, dass diese sich stark ähneln. Jedoch erhält man aus dem Verlauf von $\sigma_{\min}(\mathbf{S}_u)$ die zusätzliche Information, dass für größer werdende Winkelgeschwindigkeiten x_2 die Steuerbarkeit des Systems schlechter wird. Diese Information geht durch die Linearisierung des Systems verloren, da beim linearen System die Abhängigkeit der Steuerbarkeit von der Winkelgeschwindigkeit verschwindet. Diese zusätzliche im nichtlinearen System enthaltene Information lässt sich auch physikalisch interpretieren: Wenn sich der Stab mit einer größeren Winkelgeschwindigkeit bewegt, ist das System schwerer zu steuern. An diesem Beispiel lässt sich die Notwendigkeit für die Quantifizierung der Steuerbarkeit nichtlinearer Systeme erkennen. Jedoch lassen sich die in dieser Arbeit vorgestellten Algorithmen nicht ohne Weiteres auf nichtlineare Systeme übertragen.

Abbildung 4.14: Wagen mit Stab: $\sigma_{\min}(\tilde{\mathbf{S}}_u)$

4.2.2 Hubmagnet

Als zweites Beispiel werden die Steuerbarkeitsmaße des in Abbildung 4.15 dargestellten Hubmagneten für verschiedene Kugelmassen untersucht. Für den elektrischen Kreis gilt

$$u = Ri + \frac{d}{dt}(Li).$$

Für die Bewegung der Kugel in vertikaler Richtung

$$M \frac{d^2 y}{dt^2} = Mg + \frac{i^2}{2} \frac{dL}{dy}.$$

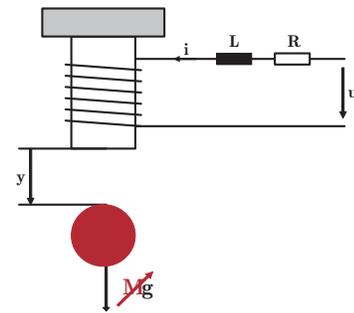


Abbildung 4.15: Hubmagnet

Die Induktivität L setzt sich aus der Selbstinduktivität L_0 und einem von der Position der Kugel abhängigen Teil zusammen. Näherungsweise gilt die Gleichung:

$$L = L_0 + L_R \frac{y_R}{y}.$$

Die Induktivität L wird umso größer, je näher sich die Eisenkugel bei dem Elektromagneten befindet. Die Induktivität L_R steht für die zusätzliche Induktivität, die durch die Eisenkugel in der Position y_R hervorgerufen wird.

Mit

$$\mathbf{x} := \left[y, \frac{dy}{dt}, i \right]^T$$

als Zustandsvektor erhält man das nichtlineare Modell

$$\begin{aligned} \frac{dx_1}{dt} &= x_2 \\ \frac{dx_2}{dt} &= g - \frac{c}{M} \frac{x_3^2}{x_1^2} \\ \frac{dx_3}{dt} &= -\frac{R}{L} x_3 + \frac{2c}{L} \frac{x_2 x_3}{x_1} + \frac{1}{L} u \end{aligned}$$

mit

$$c = \frac{L_R y_R}{2}.$$

Für die Ruhelage gilt

$$\mathbf{x}_R = \left[x_{1,R}, 0, \sqrt{\frac{Mg}{c}} x_{1,R} \right]^T \quad \text{und} \quad u_R = R \sqrt{\frac{Mg}{c}} x_{1,R}.$$

Durch Linearisierung um den Arbeitspunkt erhält man die Systemmatrizen

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 \\ \frac{2g}{y_R} & 0 & -\frac{2}{y_R} \sqrt{\frac{cg}{M}} \\ 0 & \frac{2}{L y_R} \sqrt{cMg} & -\frac{R}{L} \end{bmatrix} \quad \text{und} \quad \mathbf{b} = \begin{bmatrix} 0 \\ 0 \\ \frac{1}{L} \end{bmatrix}.$$

Im Folgenden wird die Änderung der Güte der Steuerbarkeit durch Variation der Kugelmasse m von 0,01 bis 0,1 kg betrachtet. Für die übrigen Parameter siehe Anhang B.

4.2.2.1 Modalmaße

Intuitiv kann man sich leicht vorstellen, dass bei größer werdender Masse der Kugel die Steuerbarkeit abnimmt, was auch von allen modalen Maßen angezeigt wird. Im untersuchten Intervall ändern sich die Eigenwerte der Matrix \mathbf{A} nur wenig. Sie liegen ungefähr bei $\lambda_{1,2} \approx \pm 38$ und $\lambda_3 \approx -17,5$.

Die Steuerbarkeitsmaße a nach KONNO und κ nach LITZ zeigen, dass der Eigenwert λ_3 am besten steuerbar ist. Auch die Dominanz-, Regelbarkeits- und Polempfindlichkeitsmaße liefern für diesen Eigenwert die größten Werte. Betrachtet man das Maß κ_s nach MÜLLER und LÜCKEL, so ist der Eigenwert λ_2 am besten steuerbar. Dies liegt an der Verkopplung zwischen Steuerbarkeit und Stabilität. Es ist nicht möglich, die Eigenwerte und somit die Modalmaße einer physikalischen Zustandsvariable zuzuordnen. Erst anhand des Steuerbarkeitsmaßes m nach BENNINGER kann festgestellt werden, dass die Zustandsvariable $x_3 = i$ am besten steuerbar ist. Das kann dadurch erklärt werden, dass der Strom i direkt über u beeinflusst werden kann, während die Änderung von y und $\frac{dy}{dt}$ indirekt über den Spulenstrom erfolgen muss. Als Beispiel für die Ergebnisse sind in Abbildung 4.16 das Steuerbarkeitsmaß a nach KONNO und das Steuerbarkeitsmaß m nach BENNINGER dargestellt.

4.2.2.2 Distanzmaß μ_C

Für dieses Beispiel liefern fünf der sechs Algorithmen das Ergebnis, dass korrekterweise bei größer werdender Kugelmasse das Distanzmaß μ_C kleiner wird. Wie schon bei den Beispielen zuvor liefern der Trisection-Algorithmus und die BFGS-Methode sehr enge Schranken, während die Differenz zwischen oberer und unterer Schranke bei der Bestimmung mittels der KRONECKER-Form und der Eigenwertabschätzung nach HE deutlich größer ist (vgl. Abbildung 4.17).

Die mittels KALMAN-Zerlegung ermittelte Schranke ist bis zu 24 Mal größer als die durch die übrigen Algorithmen ermittelten Schranken. Die Methode nach GAO und NEUMANN liefert auch für dieses Beispiel stückweise konstante Ergebnisse und weit auseinander liegende Schranken.

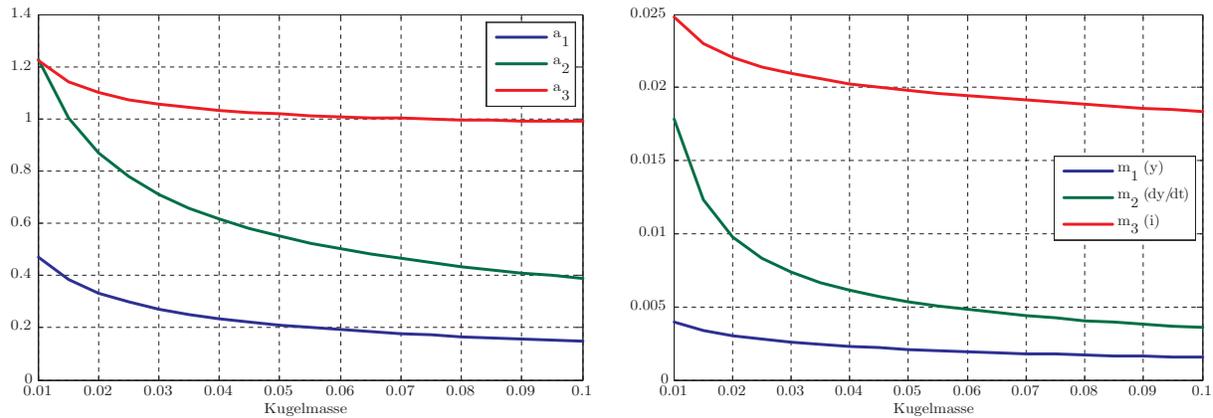


Abbildung 4.16: Steuerbarkeitsmaße a nach KONNO und m nach BENNINGER für unterschiedliche Kugelmassen

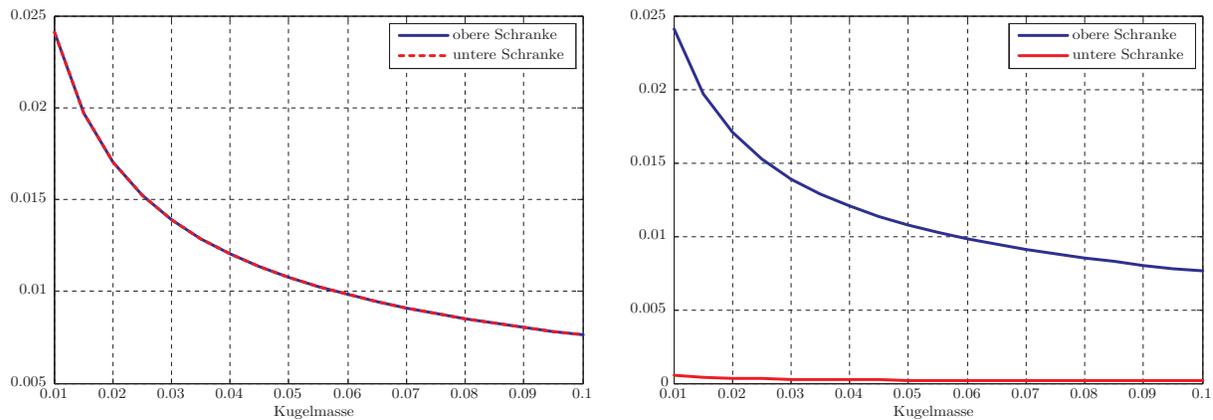


Abbildung 4.17: Obere und untere Schranken für μ_C mittels Trisection-Algorithmus (links) und KRONECKER-Form (rechts)

4.2.2.3 Distanzmaß μ_R

Durch die Betrachtung des Pseudospektrums sieht man, dass das zu bestimmende Minimum auf der reellen Achse liegt. Aus diesem Grund gilt $\mu_C = \mu_R$. Nur der Algorithmus nach GAO und NEUMANN und die Nullstellensuche von $f(s)$ liefern dieses Ergebnis. Der auf der Stufenform basierende Ansatz liefert für das gesamte Intervall $M \in [0,01 \ 0,1]$ einen konstanten Wert. Das liegt daran, dass, wenn die Matrix $[\mathbf{b}, \mathbf{A}]$ auf HESSENBERG-Form transformiert wird, das Element $h_{3,3}$ der resultierenden Matrix \mathbf{H} für jede Masse M im betrachteten Intervall gleich -1 und die übrigen Elemente betragsmäßig größer 1 sind.

Die Bestimmung von Schranken mittels Quasi-KALMAN-Zerlegung liefert Werte, die um den Faktor 10^5 größer sind als μ_C und sind damit unbrauchbar. Der Grund dafür ist auch hier, dass die Norm der Begleitmatrix \mathbf{A}_B sehr große und die beiden kleinsten Singulärwerte der Steuerbarkeitsmatrix \mathbf{S}_u sehr kleine Werte annehmen.

Die übrigen Ansätze zur Bestimmung von μ_R haben den gleichen Verlauf wie μ_C , liefern aber alle deutlich zu hohe Werte.

4.2.2.4 Einfache Methoden

Auch für dieses Anwendungsbeispiel gilt $\mu_C \approx \mu_\lambda$. Die Schätzwerte $\|A_B\| \frac{\sigma_r}{\sigma_{r-1}}$ und $\sigma_{\min}(\mathbf{S}_u)$ sind deutlich größer als das Distanzmaß μ . Die Werte für $\frac{\sigma_r}{\sigma_{r-1}}$ sind im betrachteten Intervall kleiner als μ und zeigen an, dass die Steuerbarkeit mit steigender Kugelmasse zunimmt.

4.2.3 Doppelpendel

Zwei Stäbe sind gemäß Abbildung 4.18 gelenkig und reibungsfrei auf der Masse m_3 gelagert, welche durch eine horizontale Kraft bewegt werden kann. Die Bewegung wird durch eine Feder, deren Kraft proportional zur Längenänderung wirkt, und durch viskose Reibung beeinflusst. Die Bewegungsgleichungen werden mit Hilfe der LAGRANGE-Gleichungen zweiter Art ermittelt.

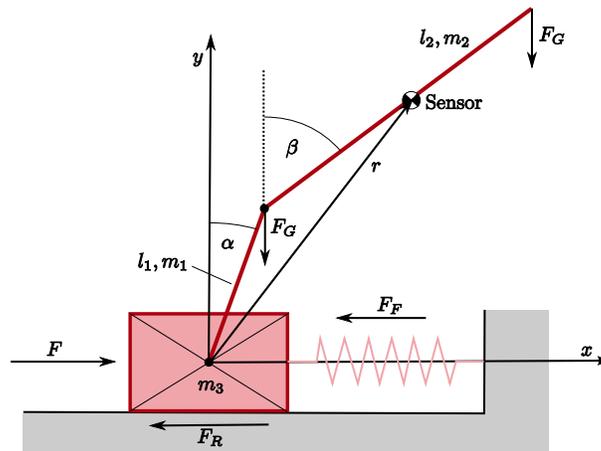


Abbildung 4.18: Doppelpendel

Das in Abbildung 4.18 dargestellte System aus drei Massenpunkten ($N = 3$) unterliegt drei Zwangsbedingungen ($k = 3$) und besitzt somit drei Freiheitsgrade ($n = 2N - k = 3$). Es sind mindestens n Ortskoordinaten zur eindeutigen Beschreibung der Lage der Massenpunkte notwendig. Durch die LAGRANGE Gleichungen 2. Art erhält man die für die Beschreibung der Bewegung der Massenpunkte notwendige minimale Anzahl von drei Gleichungen.

Generalisierte Koordinaten und Ortsvektoren: Die drei generalisierten Koordinaten ($n = 3$), die die Lage der Massenpunkte eindeutig beschreiben und die Zwangsbedingungen für beliebige Werte erfüllen, werden wie folgt gewählt:

$$\begin{aligned} q_1 &= \alpha \\ q_2 &= \beta \\ q_3 &= x. \end{aligned}$$

Dadurch ergeben sich die Ortsvektoren

$$\mathbf{r}_1 = \begin{bmatrix} q_3 + l_1 \sin q_1 \\ l_1 \cos q_1 \end{bmatrix} \quad \mathbf{r}_2 = \begin{bmatrix} q_3 + l_1 \sin q_1 + l_2 \sin q_2 \\ l_1 \cos q_1 + l_2 \cos q_2 \end{bmatrix} \quad \mathbf{r}_3 = \begin{bmatrix} q_3 \\ 0 \end{bmatrix}.$$

Kinetische Energie: Für die kinetischen Energie der drei Massenpunkte gilt

$$T_M = \sum_{i=1}^3 \frac{m_i \dot{\mathbf{r}}_i^2}{2}.$$

Da die beiden Stäbe als homogen angenommen werden, wird zusätzlich der Term

$$T_{rot} = \frac{\Theta}{2} \dot{q}_1^2 + \frac{\Theta}{2} \dot{q}_2^2$$

für die Energie der Drehbewegung der Stäbe addiert. Für das Trägheitsmoment Θ eines Stabes bezüglich einer Drehachse um einen Endpunkt gilt

$$\Theta = \frac{1}{3} m^2 l^2.$$

Für die gesamte kinetische Energie des Systems ergibt sich aus $T = T_M + T_{rot}$, somit

$$\begin{aligned} T = & \frac{1}{2} m_2 (\dot{q}_3^2 - 2\dot{q}_1 \dot{q}_3 l_1 \cos q_1 - 2\dot{q}_2 l_2 \cos q_2 [\dot{q}_3 - \dot{q}_1 l_1 \cos q_1] + 2\dot{q}_1 \dot{q}_2 l_1 l_2 \sin q_1 \sin q_2 \\ & + l_1^2 \dot{q}_1^2 + l_2^2 \dot{q}_2^2) + \frac{1}{2} m_1 (\dot{q}_3^2 - 2\dot{q}_1 \dot{q}_3 l_1 \cos q_1 + l_1^2 \dot{q}_1^2) + \frac{1}{6} m_1 l_1^2 \dot{q}_1^2 + \frac{1}{6} m_2 l_2^2 \dot{q}_2^2 + \frac{1}{2} m_3 \dot{q}_3^2. \end{aligned}$$

Potential: Die Gewichtskraft

$$\mathbf{F}_G = \begin{bmatrix} 0 \\ -mg \end{bmatrix},$$

die auf beide Stäbe wirkt, sowie die Federkraft

$$\mathbf{F}_F = \begin{bmatrix} -cq_3 \\ 0 \end{bmatrix}$$

sind konservative Kräfte und können mit Hilfe einer Potentialfunktion durch die Beziehung

$$\mathbf{F} = -\text{grad}_{\mathbf{r}_i} V$$

beschrieben werden. Für das Potential des Systems gilt

$$V = \frac{cq_3^2}{2} + m_1 g l_1 \cos q_1 + m_2 g (l_1 \cos q_1 + l_2 \cos q_2).$$

Generalisierte Kraft: Die Masse m_3 kann durch eine äußere Kraft F horizontal bewegt werden, dadurch erhält man die generalisierten Kräfte

$$Q_3 = \mathbf{F}_i^* \frac{\partial \mathbf{r}_3}{\partial q_3} = F \quad Q_1 = Q_2 = 0.$$

Dissipationfunktion: Die viskose Reibkraft

$$\mathbf{F}_R = -k\dot{\mathbf{r}}_3$$

kann mit Hilfe einer Dissipationfunktion über die Beziehung

$$\mathbf{F} = -\text{grad}_{\mathbf{r}_i} P$$

beschrieben werden. Die Dissipationfunktion lautet

$$P = \frac{k \dot{q}_3^2}{2}.$$

Bewegungsgleichungen: Durch die LAGRANGE Gleichungen 2. Art

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) - \frac{\partial L}{\partial q_i} = -\frac{\partial P}{\partial \dot{q}_i} + Q_i$$

mit $L = T - V$ erhält man drei Differentialgleichungen 2. Ordnung.

$$1: \quad \frac{1}{3} l_1 \left(3 \cos q_1 [m_2 l_2 (\ddot{q}_2 \cos q_2 - \dot{q}_2^2 \sin q_2) - \ddot{q}_3 (m_1 + m_2)] \right. \\ \left. + 3 \sin q_1 [m_2 l_2 (\dot{q}_2^2 \cos q_2 + \ddot{q}_2 \sin q_2) - g (m_1 + m_2)] + l_1 \ddot{q}_1 (3m_2 + 4m_1) \right) = 0$$

$$2: \quad \frac{1}{3} m_2 l_2 (3 \cos q_2 [l_1 (\ddot{q}_1 \cos q_1 - \dot{q}_1^2 \sin q_1) - \ddot{q}_3] + 3 \sin q_2 [l_1 (\dot{q}_1^2 \cos q_1 + \ddot{q}_1 \sin q_1) - g] \\ + 4l_2 \ddot{q}_2) = 0$$

$$3: \quad [l_1 (\dot{q}_1^2 \sin q_1 + \ddot{q}_1 \cos q_1)] (m_1 + m_2) + m_2 l_2 (\dot{q}_2^2 \sin q_2 - \ddot{q}_2 \cos q_2) + \ddot{q}_3 (m_1 + m_2 + m_3) \\ + c q_3 = F - k \dot{q}_3$$

Zustandsraummodell: Zur Beschreibung der Bewegung der drei Massenpunkte mit Hilfe eines Zustandsraummodells werden die drei Bewegungsgleichungen folgendermaßen angeschrieben:

$$\dot{\mathbf{q}} = \mathbf{v} \\ \dot{\mathbf{v}} = \mathbf{M}^{-1}(\mathbf{h} - \mathbf{p}) \quad \text{mit } \mathbf{M} = \frac{\partial \mathbf{l}}{\partial \dot{\mathbf{q}}} \quad \text{und } \mathbf{p} = \mathbf{l} - \mathbf{M}\dot{\mathbf{q}}.$$

Hierbei ist \mathbf{l} ein $(3, 1)$ -Vektor, bestehend aus den linken Seiten der Bewegungsgleichungen. Der Vektor \mathbf{h} besteht aus den rechten Seiten der Gleichungen. Der Vektor \mathbf{v} ist somit nur von \mathbf{q} und \mathbf{v} bzw. $\dot{\mathbf{q}}$ abhängig, aber nicht von $\ddot{\mathbf{q}}$.

Linearisierung: Durch Einführung des Vektors

$$\mathbf{x} := \begin{bmatrix} \mathbf{q} \\ \mathbf{v} \end{bmatrix} = \left[\alpha \quad \beta \quad x \quad \frac{d\alpha}{dt} \quad \frac{d\beta}{dt} \quad \frac{dx}{dt} \right]^T$$

ergibt sich ein nichtlineares System der Form $\frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x}, F)$. Für $F = 0$ erhält man für die Ruhelagen des Systems

$$\mathbf{f}(\mathbf{x}_R) = \mathbf{0} \quad \curvearrowright \quad \mathbf{x}_R = [0, 0, 0, l\pi, k\pi, 0]^T \quad \text{mit } k, l \in \mathbb{Z}.$$

Durch Linearisierung um die instabile Ruhelage $\mathbf{x}_R = \mathbf{0}$, d.h., beide Stäbe stehen senkrecht nach oben, und durch Einführen der Konstanten

$$M_1 = 3m_2m_3 + m_1m_2 + 4m_1^2 + 16m_1m_3,$$

$$M_2 = 5m_1m_2 + 4m_1^2 + m_2^2,$$

$$M_3 = 3m_2m_3 + 4m_1m_3 + m_1m_2 + m_1^2$$

und

$$M_4 = 3m_2m_3 + m_1m_2 + 4m_1^2 + 16m_1m_3$$

können die Systemmatrizen wie folgt angeschrieben werden:

$$\mathbf{A} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ \frac{3g(M_2+4m_2m_3+4m_1m_3)}{l_1M_1} & -\frac{9m_3m_2g}{l_1M_1} & \frac{3c(4m_1+m_2)}{l_1M_1} & 0 & 0 & \frac{3k(4m_1+m_2)}{l_1M_1} \\ -\frac{9m_3g(m_2+m_1)}{l_2M_1} & \frac{3gM_3}{l_2M_1} & \frac{3m_1c}{l_2M_1} & 0 & 0 & \frac{3m_1k}{l_2M_1} \\ -\frac{3gM_2}{M_1} & -\frac{3m_1m_2g}{M_1} & -\frac{c(16m_1+3m_2)}{M_1} & 0 & 0 & -\frac{k(16m_1+3m_2)}{M_1} \end{bmatrix}$$

$$\mathbf{b} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ -\frac{3(m_2+4m_1)}{l_1M_4} \\ -\frac{3m_1}{l_2M_4} \\ \frac{16m_1+3m_2}{M_4} \end{bmatrix}.$$

Die Zahlenwerte der Parameter sind im Anhang B angegeben.

Mit Hilfe dieses Beispiels wird gezeigt, dass die in dieser Arbeit erhaltenen Erkenntnisse bezüglich der Steuerbarkeit einfach auf die Beobachtbarkeit übertragbar sind. Aussagen über die Beobachtbarkeit liefern auf Grund der Dualität die Steuerbarkeitsmaße von $(\mathbf{A}^T, \mathbf{C}^T)$. Im Folgenden werden die Beobachtbarkeitsmaße für verschiedene „Sensorpositionen“ berechnet. Hierfür wird der vektorielle Ausgang gleich den Ortskoordinaten der Sensorposition gewählt (vgl. Abbildung 4.18):

$$\mathbf{r} = \mathbf{y} = \begin{bmatrix} \alpha \sin q_1 - (\alpha - l_1)\sigma(\alpha - l_1) \sin q_1 + (\alpha - l_1)\sigma(\alpha - l_1) \sin q_2 + q_3 \\ \alpha \cos q_1 - (\alpha - l_1)\sigma(\alpha - l_1) \cos q_1 + (\alpha - l_1)\sigma(\alpha - l_1) \cos q_2 \end{bmatrix}$$

Der Parameter α entspricht dem auf den Stäben zurückgelegten Weg. Wird α von Null bis Sieben variiert, so wandert der Sensor den Stäben entlang bis zur Spitze des zweiten Stabes. Der Ausgang wird um die Ruhelage \mathbf{x}_R linearisiert. Somit erhält man für die linearisierte Ausgangsmatrix

$$\mathbf{C} = \begin{bmatrix} \alpha - (\alpha - l_1)\sigma(\alpha - l_1) & (\alpha - l_1)\sigma(\alpha - l_1) & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

4.2.3.1 Modalmaße

Für dieses Beispiel kann eine physikalische Interpretation folgendermaßen aussehen (vgl. Abbildung 4.19):

Die Maße κ_1 und κ_2 beschreiben die Beobachtbarkeit des Stabes 1 mit der Länge $l_1 = 2$. Das Maß nimmt bis zur Spitze des Stabes bei $\alpha = 2$ zu und nimmt dann wieder ab, je weiter sich

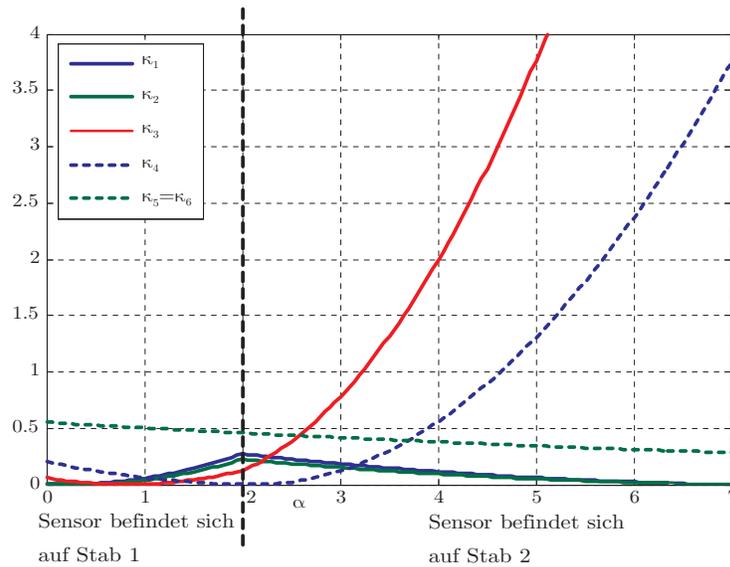


Abbildung 4.19: Beobachtbarkeitsmaße κ nach LITZ für unterschiedliche Sensorpositionen

der Sensor von Stab 1 entfernt. Die Maße κ_3 und κ_4 messen die Beobachtbarkeit des 2. Stabes. Auch hier nimmt das Maß zu, je weiter sich der Sensor der Spitze des Stabes nähert. Die Maße geben somit an, dass je weiter sich der Sensor an der Stabspitze befindet, desto größer ist die Beobachtbarkeit. Dies lässt sich dadurch erklären, dass die Spitze mit größerer Geschwindigkeit längere Wege durchläuft als Punkte im Stabinneren, weshalb die Bewegung besser zu beobachten ist.

Das Beobachtbarkeitsmaß m nach BENNINGER liefert für dieses Beispiel keine zufrieden stellenden Ergebnisse. Der Grund dafür sind numerische Probleme bei der Invertierung der GRAM-schen Matrix \mathbf{Q}_S , da diese fast singularär ist. Dies ist durch den kleinsten Singulärwert der Matrix \mathbf{Q}_S erkennbar, der für verschiedene Sensorpositionen kleiner 10^{-15} ist.

4.2.3.2 Distanzmaß μ_C

In Abbildung 4.20 und 4.21 sind beispielhaft die obere und untere Schranke von μ_C dargestellt, die durch den auf der Eigenwertabschätzung nach HE basierenden Ansatz bzw. durch den Trisection-Algorithmus ermittelt wurden. Die Maße zeigen, dass die Beobachtbarkeit schlecht ist, wenn sich der Sensor am Ankerpunkt des 2. Stabes befindet. Weiters wird angezeigt, dass die bestmögliche Beobachtbarkeit erreicht wird, wenn sich der Sensor ungefähr bei einem Drittel des 2. Stabes befindet. Fraglich ist wie dieser Sachverhalt physikalisch zu interpretieren ist, da ein „Knick“ der Beobachtbarkeit, während sich der Sensor auf Stab 2 befindet, unlogisch erscheint.

Der Ansatz nach HE liefert für dieses Beispiel eine sehr kleine untere Schranke, wodurch der Bereich zwischen oberer und unterer Schranke, in dem sich μ_C befinden kann, groß ist.

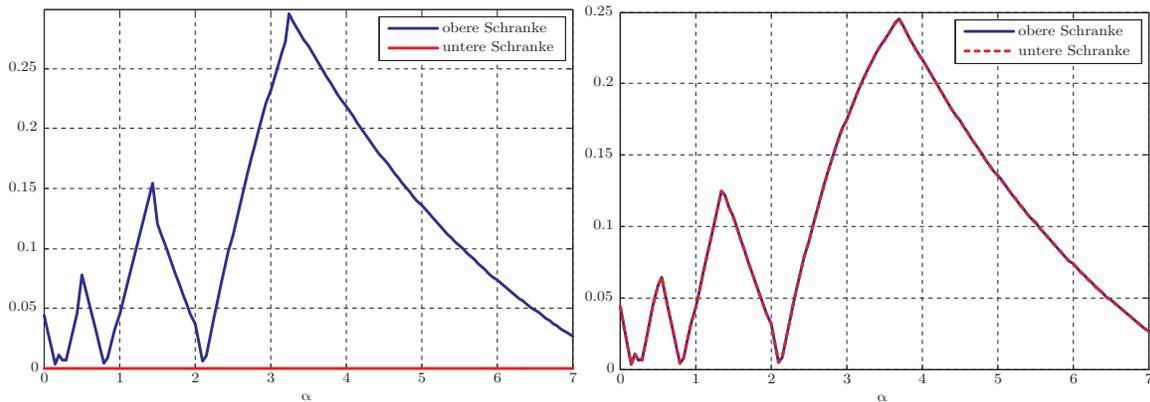


Abbildung 4.20: Schranken für μ_C mittels Eigenwertabschätzung nach HE

Abbildung 4.21: Schranken für μ_C , ermittelt durch den Trisection-Algorithmus

4.2.3.3 Distanzmaß μ_R

Auch für dieses Beispiel gilt $\mu_R = \mu_C$. Abbildung 4.22 zeigt das durch den Algorithmus von GAO und NEUMANN erhaltene Ergebnis, welches dem durch die Minimumsuche von $f(s)$ entspricht. Dies sind die einzigen der fünf vorgestellten Methoden, deren Ergebnisse mit dem für μ_C bestimmten Schranken übereinstimmen. Die übrigen Methoden liefern obere Schranken, die den Verlauf, mit Ausnahme der Schranke nach GAHINET, nicht wiedergeben (vgl. Abbildung 4.22).

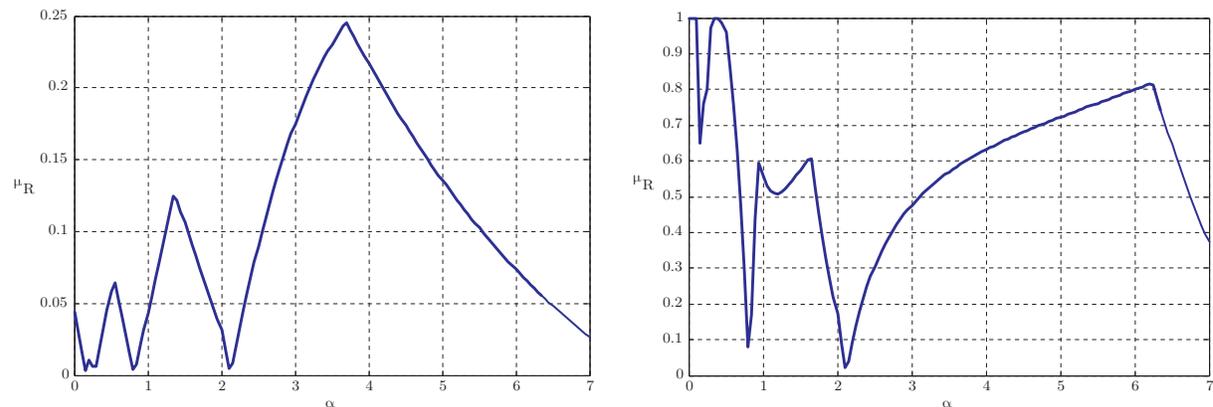


Abbildung 4.22: Ergebnisse der Bestimmung von μ_R mittels des Algorithmus nach GAO und NEUMANN und der Nullstellensuche von $f(s)$ (links) und mittels Stufenform (rechts)

4.2.3.4 Einfache Methoden

Die obere Schranke μ_λ liefert einen sehr ähnlichen Verlauf, wie das durch den Trisection-Algorithmus ermittelte Distanzmaß μ_C . Der einfach zu berechnenden Schätzwert $\sigma_{\min}(\mathbf{S}_u)$ liefert Werte in der Größenordnung von μ_C , gibt aber den Verlauf von μ_C nicht korrekt wieder. Auch in diesem Beispiel liefert die Abschätzung $\|\mathbf{A}_B\| \frac{\sigma_r}{\sigma_{r-1}}$ Werte, die deutlich größer sind als μ_C .

4.2.4 Zusammenfassung der Beispiele

Um einen Überblick zu erhalten, sind die Ergebnisse der verschiedenen Maßzahlen für die durchgeführten Beispiele in nachfolgender Tabelle stichwortartig zusammengefasst.

Die in diesem Kapitel angeführten Beispiele zeigen, dass die Modalmaße, die Energiemaße und auch das Distanzmaß in einigen Fällen gute Ergebnisse liefern, die sich auch physikalisch interpretieren lassen.

Die Beispiele haben folgende Tatsachen bestätigt: Das Maß κ_S nach MÜLLER und LÜCKEL liefert wegen der Verkopplung von Stabilität und Steuerbarkeit oft unbrauchbare Ergebnisse. Die Polempfindlichkeitsmaße nach LITZ werden auf Grund der großen Anzahl an Maßzahlen sehr schnell unübersichtlich. Die übrigen Modalmaße liefern sehr ähnliche Ergebnisse. Das Beispiel „Doppelpendel“ zeigt, dass obwohl diese Maße Aussagen im Modalraum treffen, die Ergebnisse in manchen Fällen physikalisch interpretierbar sind.

Obwohl die Probleme bei der Berechnung der GRAMschen Matrix bei mehrfachen Eigenwerten durch den Reihenansatz umgangen werden können, bleiben die numerischen Probleme bei der Inversion dieser Matrix bestehen, was zu fehlerhaften Ergebnissen führen kann.

Über die Berechnung des Distanzmaßes wird Folgendes festgehalten: Für die Bestimmung von μ_C liefern der Trisection-Algorithmus sowie die BFGS-Methode die besten und engsten Schranken. Auch die Schätzung von μ_C nach WICKS und DeCARLO liefert, bei geeignet gewähltem Startwert, gute Ergebnisse. Der Ansatz über die KRONECKER-Form liefert eine gute obere Schranke, jedoch ist die untere Schranke in vielen Fällen viel kleiner als der wahre Wert von μ_C . Analoges gilt bei Systemen mit einfachen Eigenwerten auch für den Ansatz über die Eigenwertabschätzung nach HE. Die durch die übrigen Methoden ermittelten Werte sind i.A. viel größer als μ_C . Für die Bestimmung von μ_R liefern der Algorithmus nach GAO und NEUMANN sowie die Nullstellensuche von $f(s)$ die besten und engsten Schranken. Auch hier liefern die übrigen Methoden meist Werte, die um ein Vielfaches größer sind als der „wahre“ Wert von μ_R .

	Beispiel 2.2	Beispiel 2.3	Wagen mit Stab	
Pseudospektrum	zwei lokale Minima auf der reellen Achse ($\Rightarrow \mu_C = \mu_R$)	ein Minimum auf der reellen Achse ($\Rightarrow \mu_C = \mu_R$)	im Intervall $\alpha \in [-\pi/2, \pi/2]$ zwei Minima auf der reellen Achse ($\Rightarrow \mu_C = \mu_R$); im Intervall $\alpha \in [\pi/2, -\pi/2]$ ein konjugiert komplexes Minimum ($\Rightarrow \mu_C \neq \mu_R$)	
Modalmäße	KONNO a (Gl. 2.2)	Unstetigkeitsstelle: $\alpha = -2$ (doppelter EW)	✓	
	LÜCKEL, MÜLLER κ_s (Gl. 2.3)	✓	inkonsistent (doppelter EW)	
	LITZ κ (Gl. 2.4)	✓	inkonsistent (doppelter EW)	
	abs. Polempf. S (Gl. 2.6)	Unstetigkeitsstelle: $\alpha = -2$	inkonsistent, konstant	✓
	rel. Polempf. s (Gl. 2.7)	Unstetigkeitsstellen: $\alpha = -2$ (doppelter EW) $\alpha = 0$ (EW bei 0)	inkonsistent, konstant	✓
	Maximalmaß M (Gl. 2.8)	Unstetigkeitsstellen: $\alpha = -2, \alpha = 0$	inkonsistent	✓
	Summenmaß Σ (Gl. 2.9)	Unstetigkeitsstellen: $\alpha = -2, \alpha = 0$	inkonsistent	✓
	BENNINGER m (Gl. 2.14)	✓	✓	✓
	BENNINGER v (Gl. 2.15)	✓	✓	✓
	μ_C	Trisection mit GUs Test	✓	✓ (für $\alpha \pm \frac{\pi}{2}$ nicht berechenbar)
BFGS mit GUs Test		✓	✓ (für $\alpha \pm \frac{\pi}{2}$ nicht berechenbar)	
KRONECKER-Form (KCF)		gute obere Schranke	gute obere Schranke	
Eigenwertabschätzung nach HE		gute obere Schranke; untere Schranke ist 0 bei $\alpha = -2$	gute obere Schranke (untere Schranke besser als mit KCF)	
WICKS und DeCARLO		für $\alpha \in [0, 1]$ nur lokales Minimum gefunden	✓	✓ (für $\alpha \pm \frac{\pi}{2}$ nicht berechenbar)
GAO und NEUMANN		stückweise konstant	untere Schranke größer μ_C ; inkonsistent	stückweise konstant; weit auseinander liegende Schranken
μ_R	GAHINET (Gl. 3.28)	bis zu Faktor 3 zu groß	bis zu Faktor 6 zu groß	
	Nullstellen von $f(s)$	✓	✓	
	Quasi-KALMAN-Zerlegung	bis zu Faktor 3 zu groß	bis zu Faktor 2 zu groß	fälschlicherweise hohe Werte in Umgebung der Unstetigkeitsstellen $\alpha = \pm\pi/2$
	Stufenform	bis zu Faktor 3 zu groß	✓	konstant
	GAO und NEUMANN	✓	✓	✓

	Hubmagnet	Doppelpendel	Besonderheiten	
Pseudospektrum	3 lokale Minima auf der reellen Achse ($\Rightarrow \mu_C = \mu_R$)	3 konjugiert komplexe Minima ($\Rightarrow \mu_C \neq \mu_R$)		
Modalmaße	KONNO a (Gl. 2.2)	✓		
	LÜCKEL, MÜLLER κ_s (Gl. 2.3)	bewertet als einziges Maß $\lambda_2 = -38$ als am besten steuerbar	Verzerrung auf Grund der Verkopplung mit Stabilität	
	LITZ κ (Gl. 2.4)	✓		
	abs. Polempf. S (Gl. 2.6)	✓	liefert komplexe Werte	
	rel. Polempf. s (Gl. 2.7)	✓	liefert komplexe Werte	
	Maximalmaß M (Gl. 2.8)	✓		
	Summenmaß Σ (Gl. 2.9)	✓		
	BENNINGER m (Gl. 2.14)	✓	keine brauchbaren Ergebnisse (da GRAMsche Matrix „nahezu singular“)	konsistent bei Steuerbarkeitsverlust durch innere Verkopplungen
	BENNINGER v (Gl. 2.15)	✓		
μ_C	Trisection mit GUs Test	✓	enge Schranken abhängig von tol	
	BFGS mit GUs Test	✓	✓ (Knick bei $\alpha = 2,75$)	enge Schranken abh. von tol ; falsches Ergebnis bei ungünstigem Startwert
	KRONECKER-Form (KCF)	gute obere Schranke	gute obere Schranke (Knick bei $\alpha = 2,75$)	weit auseinanderliegende Schranken (keine Toleranz einstellbar); „Rauschen“; untere Schranke oft sehr klein; schneller Algorithmus
	Eigenwertabschätzung nach HE	gute obere Schranke; untere besser als mit KCF	untere Schranke schlechter als mit KCF (Knick bei $\alpha = 2,75$)	gute obere Schranke; untere oft sehr klein; schneller Algorithmus
	WICKS und DeCARLO	findet nur lokales Minimum	✓ (Knick bei $\alpha = 2,45$)	schneller Algorithmus; findet manchmal nur lokales Minimum
	GAO und NEUMANN	stückweise konstant	Schranken um Faktor 2 zu groß; stückweise konstant	liefert wegen fest vorgegebenen Gitters stückweise konstante Ergebnisse
	GAHINET (Gl. 3.28)	bis zu Faktor 76 zu hoch	um Faktor 6 zu groß	hohe Werte; abhängig davon wie gut μ_c bestimmt wurde
	Nullstellen von $f(s)$	bis zu Faktor 2,4 zu hoch	Verlauf gleicht dem durch Trisection ermittelten μ_c	schneller & zuverlässiger Algorithmus
	Quasi-KALMAN-Zerlegung	um $\approx 10^5$ zu hoch	um Faktor 30 größer als bei Nullstellensuche von $f(s)$	liefert i.A. sehr hohe Werte
	Stufenform	konstant	Schranken sehr viel größer als μ_c	wenn System bereits in Stufenform vorliegt, ist Ergebnis konstant
GAO und NEUMANN	✓	(Knick bei $\alpha = 2$)	schneller & zuverlässiger Algorithmus	
μ_R				

4.3 Anwendungen

Zum Abschluss dieses Kapitels werden Anwendungen der Steuerbarkeitsmaße bei der Analyse von Systemen bzw. bei der Synthese von Regelkreisen vorgestellt.

Die Wahl des Steuerbarkeitsmaßes ergibt sich direkt aus der jeweiligen Anwendung. Für diese ist entscheidend, ob die Steuerbarkeit des Gesamtsystems oder der einzelnen Zustandsvariablen bewertet werden soll.

Betrachtet man z.B. das Problem der Platzierung von Aktoren und Sensoren, eignet sich das Energiemaß m nach BENNINGER besonders gut, da einzelne Zustandsvariablen bewertet werden. Ein Beispiel für die Stellgliedpositionierung einer flexiblen Raumfahrtstruktur mit Hilfe eines Energiemaßes ist in [1] aufgeführt.

Da Modalmaße die Eigenwerte bewerten, sind diese für die Ordnungsreduktion gut geeignet. Die Ordnung eines komplexen Systems kann durch Vernachlässigung der Systemteile mit geringem Einfluss auf das Gesamtsystem reduziert werden. Die Modellreduktion mit Hilfe von Modalmaßen wurde erfolgreich für unterschiedlichste Systeme angewendet, wie z.B. für einen Dampferzeuger, in [1], eine Destillationskolonne, in [29], eine Durchlaufglühanlage, in [31] und für einen Hinterachsprüfstand, in [13].

Wie am Anwendungsbeispiel „Doppelpendel“ gezeigt, können die Maße bei der Wahl von Aktor- und Sensorpositionen nützlich sein. Weitere Beispiele dafür können für die Sensorplatzierung an einem elastischen Roboter in [39] nachgelesen werden.

Das Distanzmaß, welches die Steuerbarkeit des Gesamtsystems bewertet, kann folgendermaßen eingesetzt werden: Da die numerische Bestimmung der Steuerbarkeit mit Hilfe der Kriterien nach KALMAN oder HAUTUS in manchen Fällen zu fehlerhaften Ergebnissen führt, wird das Distanzmaß μ für eine zuverlässige Bestimmung der Steuerbarkeit benutzt. PAIGE führt in [38] eine Schranke μ_0 ein, welche sich aus Unsicherheiten der Systemmatrizen und aus einer Schranke für die zu erwartenden numerischen Fehler bei der Berechnung von μ zusammensetzt. Ist das berechnete Distanzmaß μ größer als der Wert μ_0 , dann ist das System mit Sicherheit steuerbar. Anderenfalls kann keine zuverlässige Aussage über die Steuerbarkeit des Systems gemacht werden.

Weiters kann eine Aussage über die Empfindlichkeit des Systems auf Grund von Parameterrungenauigkeiten getroffen werden: Wenn das (nicht exakt) modellierte System beispielsweise schlecht steuerbar ist, ist zu erwarten, dass das reale System auf Grund von Parametertoleranzen oder -schwankungen nicht steuerbar ist.

Auch für die Wahl bzw. die Bewertung eines Arbeitspunktes eines nichtlinearen Systems können die Maße hilfreich sein. Wird ein nichtlineares System

$$\frac{dx}{dt} = f(x, u)$$

um einen Arbeitspunkt (x_R, u_R) linearisiert, um einen linearen Zustandsregler zu entwerfen, so entsteht ein lineares System, dessen Steuerbarkeitsgüte mit Hilfe der Maße bewertet werden kann. Wird ein guter Arbeitspunkt dadurch definiert, dass das linearisierte System gut steuerbar ist, kann somit mittels der Maße die Qualität des Arbeitspunktes bewertet werden.

Im Weiteren wird auf drei Anwendungsgebiete näher eingegangen. In Abschnitt 4.3.1 wird untersucht, ob die Maßzahlen bei der Wahl der Eigenwerte für den Zustandsreglerentwurf nütz-

lich sind. In Abschnitt 4.3.2 wird die Aussagekraft der Maßzahlen bei der Wahl der Sensorpositionierung zur Fehlerdiagnose behandelt. D.h., es wird untersucht, wie mit Hilfe der Maßzahlen Sensoren zu platzieren sind, dass auftretende Fehler detektiert und isoliert werden können. In Abschnitt 4.3.3 wird die Optimierung der Systemdaten, insbesondere des Eingangsvektors, bezüglich des Steuerbarkeitsmaßes behandelt.

4.3.1 Zustandsreglerentwurf

Für ein steuerbares LZI-System

$$\frac{d\mathbf{x}}{dt} = \mathbf{A}\mathbf{x} + \mathbf{b}u$$

wird ein Zustandsregler

$$u = -\mathbf{h}^T \mathbf{x} + Vr$$

entworfen. Für den geschlossenen Regelkreis ergibt sich die Zustandsbeschreibung

$$\frac{d\mathbf{x}}{dt} = \tilde{\mathbf{A}}\mathbf{x} + \tilde{\mathbf{b}}r$$

mit

$$\tilde{\mathbf{A}} = \mathbf{A} - \mathbf{b}\mathbf{h}^T \quad \text{und} \quad \tilde{\mathbf{b}} = \mathbf{b}V.$$

Damit der Regler für konstante Eingangsgrößen stationär genau ist, wird der Verstärkungsfaktor V wie folgt gewählt

$$V = -\frac{1}{\mathbf{c}^T(\mathbf{A} - \mathbf{b}\mathbf{h}^T)^{-1}\mathbf{b}}. \quad (4.2)$$

Durch Vorgabe der gewünschten n Eigenwerte $\tilde{\lambda}_i$ ($i = 1, \dots, n$) der Matrix $\tilde{\mathbf{A}}$ des geschlossenen Regelkreises kann der Vektor \mathbf{h}^T bestimmt werden.

In [34] wird die Konditionierung des Problems der Polvorgabe beim Zustandsreglerentwurf in Abhängigkeit von Ungenauigkeiten bzw. Störungen der Systemdaten analysiert. Hierbei zeigt sich, dass ein Zusammenhang zwischen der Konditionierung und der Konditionszahl der Eigenvektormatrix des geschlossenen Kreises $\tilde{\mathbf{A}}$, der Norm des Rückkopplungsvektors \mathbf{h} und des Distanzmaßes μ_C besteht. Arbeitet man mit gestörten Systemdaten, kann ein kleines Distanzmaß μ_C zu einem großen Fehler zwischen dem Rückführvektor des gestörten und des nichtgestörten Systems führen. Ist μ_C klein, kann somit nicht davon ausgegangen werden, dass die Wunscheigenwerte mit den Eigenwerten des geschlossenen Kreises übereinstimmen.

Im Gegensatz dazu wird hier die Frage beantwortet, wie die Wunscheigenwerte plziert werden müssen um einen möglichst großen Wert für das Distanzmaß μ_C zu erhalten. Die Wahl der Wunscheigenwerte $\tilde{\lambda}_i$ ist i.A. unter Berücksichtigung der Stabilität frei. Durch die Vorgabe eines gewünschten Reglerverhaltens (maximales Überschwingen, Mindest-Anstiegszeit, Stellgrößenbeschränkung, ...), wird die Position der Eigenwerte λ auf einen Bereich der s-Ebene eingeschränkt. Es existieren i.A. keine Methoden zur Bestimmung der Wunscheigenwerte $\tilde{\lambda}_i$. Mit Hilfe der Steuerbarkeitsmaße können die Eigenwerte so innerhalb dieses Bereiches platziert werden, dass die Steuerbarkeit des geregelten Systems maximal wird, was im Folgenden für Systeme 1. und 2. Ordnung gezeigt wird. In Reglerkaskaden kann durch diese Eigenwertplatzierung die Steuerbarkeit des inneren Kreis maximiert werden.

4.3.1.1 Systeme 1. Ordnung

Für ein System 1. Ordnung

$$\frac{dx}{dt} = ax + bu \quad y = cx$$

erhält man für den geschlossenen Regelkreis

$$\frac{dx}{dt} = (a - bh)x + bVr = \tilde{\lambda}x + \tilde{b}r.$$

Der Reglerparameter h ergibt sich durch Vorgabe des Wunscheigenwertes $\tilde{\lambda}$ zu

$$\tilde{\lambda} = a - bh \quad \Leftrightarrow \quad h = -\frac{\tilde{\lambda} - a}{b}.$$

Für den Verstärkungsfaktor gilt

$$V = -\frac{\tilde{\lambda}}{cb}.$$

Damit erhält man das geregelte System

$$\frac{dx}{dt} = \tilde{\lambda}x - \frac{\tilde{\lambda}}{c}r.$$

Man kommt zu der Schlussfolgerung, dass wenn $|\tilde{b}| = \left| -\frac{\tilde{\lambda}}{c} \right|$ größer wird, die Steuerbarkeit zunimmt. D.h., je weiter links der Wunscheigenwert $\tilde{\lambda}$ in der s -Ebene liegt, desto besser steuerbar ist das System.

Die Maße, die durch den Eigenwert normiert werden (relatives Polempfindlichkeitsmaß $s = -1$ nach LITZ, Summen- und Maximalmaß $M = 1$ und $\Sigma = 1$ nach LITZ) zeigen an, dass die Steuerbarkeit des geregelten Systems unabhängig von der Wahl von $\tilde{\lambda}$ konstant bleibt. Die übrigen Maße (Steuerbarkeitsmaß $\kappa = \frac{\tilde{\lambda}^2}{c^2}$, absolutes Polempfindlichkeitsmaß $S = -\tilde{\lambda}$ nach LITZ und Distanzmaß $\mu = \left| \frac{-\tilde{\lambda}}{c} \right|, \dots$) zeigen, dass die Steuerbarkeit mit kleiner werdenden Werten für $\tilde{\lambda}$ ($\tilde{\lambda} < 0$) zunimmt.

Betrachtet man für das nicht geregelte System das Steuerbarkeitsmaß $\kappa = b^2$ und das Distanzmaß $\mu = |b|$, so kommt man zu dem Ergebnis, dass für $\tilde{\lambda} < -|cb|$ die Steuerbarkeit des geregelten Systems größer ist als die des nicht geregelten Systems.

4.3.1.2 Systeme 2. Ordnung

Gegeben sei das System 2. Ordnung

$$\begin{aligned} \frac{dx}{dt} &= \mathbf{A}\mathbf{x} + \mathbf{b}u = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \mathbf{x} + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} u \\ y &= \mathbf{c}^T \mathbf{x} = [c_1 \ c_2] \mathbf{x}. \end{aligned}$$

Für eine übersichtlichere Darstellung der nachfolgenden Ergebnisse werden die Abkürzungen

$$\alpha := a_{12}b_2 - a_{22}b_1 \quad \text{und} \quad \beta := a_{21}b_1 - a_{11}b_2 \quad (4.3)$$

eingeführt. Für die Wunscheigenwerte $\tilde{\lambda}_{1,2}$ erhält man für den Zustandsregler den Vektor

$$\mathbf{h} = \begin{bmatrix} \frac{\beta(a_{11}-\tilde{\lambda}_2-\tilde{\lambda}_1)-a_{21}\alpha-b_2\tilde{\lambda}_1\tilde{\lambda}_2}{b_1\beta-b_2\alpha} \\ \frac{\alpha(-a_{22}+\tilde{\lambda}_1+\tilde{\lambda}_2)+a_{12}\beta+b_1\tilde{\lambda}_1\tilde{\lambda}_2}{b_1\beta-b_2\alpha} \end{bmatrix}$$

und für den Verstärkungsfaktor gilt

$$V = -\frac{\tilde{\lambda}_1\tilde{\lambda}_2}{c_1\alpha + c_2\beta}.$$

Um die Modalmaße zu berechnen, müssen für das geregelte System

$$\begin{aligned} \frac{d\mathbf{x}}{dt} &= (\mathbf{A} - \mathbf{b}\mathbf{h}^T)\mathbf{x} + \mathbf{b}Vr = \tilde{\mathbf{A}}\mathbf{x} + \tilde{\mathbf{b}}r \\ y &= \mathbf{c}^T\mathbf{x} \end{aligned}$$

die Eigenvektoren bestimmt werden. Eine mögliche Wahl für die Rechts-Eigenvektoren ist

$$\mathbf{X}_R = \begin{bmatrix} \frac{\alpha+b_1\tilde{\lambda}_1}{\beta+b_2\tilde{\lambda}_1} & \frac{\alpha+b_1\tilde{\lambda}_2}{\beta+b_2\tilde{\lambda}_2} \\ 1 & 1 \end{bmatrix}.$$

Damit kann nun das Steuerbarkeitsmaß κ nach LITZ

$$\kappa_i = \frac{\mathbf{x}_{Li}^H \tilde{\mathbf{b}}\tilde{\mathbf{b}}^T \mathbf{x}_{Li}}{\mathbf{x}_{Li}^H \mathbf{x}_{Li}} = V^2 \frac{\mathbf{x}_{Li}^H \mathbf{b}\mathbf{b}^T \mathbf{x}_{Li}}{\mathbf{x}_{Li}^H \mathbf{x}_{Li}}$$

berechnet werden. Mit dem Verstärkungsfaktor V aus Gleichung (4.2), den Links-Eigenvektoren \mathbf{x}_{Li} , die den Zeilen der Matrix $\mathbf{X}_L = \mathbf{X}_R^{-1}$ entsprechen, und den Abkürzungen α und β gemäß (4.3) erhält man das Steuerbarkeitsmaß

$$\kappa_{1,2} = \frac{(b_1\beta - b_2\alpha)^2(\tilde{\lambda}_{1,2}\tilde{\lambda}_{2,1})^2}{(c_1\alpha + c_2\beta)^2(\beta^2 + \alpha^2 + 2\text{Re}\{\tilde{\lambda}_{2,1}\}(b_2\beta + b_1\alpha) + (b_1^2 + b_2^2)|\tilde{\lambda}_{2,1}|^2)}.$$

Für $\tilde{\lambda}_{1,2} \rightarrow -\infty$ streben die Maße gegen Unendlich. Somit gilt, wie schon bei Systemen 1. Ordnung, dass je kleiner die Wunscheigenwerte $\tilde{\lambda}_{1,2}$ gewählt werden, desto steuerbarer wird das geregelte System.

Berechnet man für das geregelte System das Polempfindlichkeitsmaß, so erhält man

$$S_{1,2} = \frac{\tilde{\lambda}_{1,2}\tilde{\lambda}_{2,1}(c_1\alpha + c_2\beta + c_1b_1\tilde{\lambda}_{1,2} + c_2b_2\tilde{\lambda}_{1,2})}{(\tilde{\lambda}_{1,2} - \tilde{\lambda}_{2,1})(c_1\alpha + c_2\beta)}.$$

Für $\tilde{\lambda}_1 \rightarrow -\infty$ bzw. $\tilde{\lambda}_2 \rightarrow -\infty$ streben die Maße S_1 bzw. S_2 gegen $\pm\infty$. Dieses Resultat stimmt mit den aus dem Steuerbarkeitsmaß κ nach LITZ erhaltenen Ergebnis überein. Für $\tilde{\lambda}_1 = \tilde{\lambda}_2$ besitzt das Polempfindlichkeitsmaß eine Unstetigkeitsstelle, in deren Umgebung der Term $\tilde{\lambda}_{1,2} - \tilde{\lambda}_{2,1}$ im Nenner zu hohen Werten führt, wodurch für nahe beieinander liegende Eigenwerte $\lambda_{1,2}$ eine gute Steuerbarkeit angezeigt wird.

Das Distanzmaß μ ist für diesen Fall nicht analytisch berechenbar. Deshalb wurde, um die Auswirkungen der Wunscheigenwerte auf das Maß zu untersuchen, die numerische Lösung herangezogen. Hierfür wurden die Systemdaten mit unterschiedlichen Zufallszahlen gefüllt. Die Simulationen zeigen, dass das Distanzmaß μ für weit links liegende Wunscheigenwerte $\tilde{\lambda}_{1,2}$ i.A. recht hohe Werte liefert. Im Gegensatz zu den Modalmaßen erhält man nicht unabhängig von den Systemdaten die größten Maßwerte, wenn die Wunscheigenwerte gegen Unendlich streben. Das Distanzmaß kann durchaus das Maximum für Wunscheigenwerte, die nahe bei der imaginären Achse liegen, erreichen.

Die in diesem Abschnitt durchgeführten Untersuchungen zeigen, dass das Steuerbarkeitsmaß des zustandsgeregelten Systems für kleine Wunscheigenwerte $\tilde{\lambda}$ i.A. groß ist. Somit sollten die Wunscheigenwerte innerhalb des oben erwähnten Bereichs möglichst weit „links“ liegen.

Für Systeme höherer Ordnung ist eine analytische Berechnung nicht mehr sinnvoll bzw. nicht möglich. In [40] wird anhand zweier ausgewählter Systeme 3. Ordnung gezeigt, dass das Distanzmaß i.A. größer wird, je weiter links die Eigenwerte liegen.

4.3.2 Sensorplatzierung zur Fehlerdiagnose

In diesem Abschnitt wird gezeigt, wie das Steuerbarkeitsmaß nach BENNINGER bei der Fehlerdiagnose eingesetzt werden kann. Mit Hilfe des Maßes wird eine Aussage darüber getroffen, welcher Fehler auf welche Zustandsvariable wirkt. Daraus kann ermittelt werden, welche Zustandsvariablen für die Fehlerdiagnose gemessen werden müssen. Vorab werden benötigte Begriffe (vgl. z.B. [33]) erläutert und einige Annahmen getroffen:

Fehler Ein Fehler ist eine unzulässige Abweichung eines oder mehrerer Merkmale einer technischen Anlage. Diese äußern sich in einer Veränderung der gemessenen Signale. In der hier durchgeführten Untersuchung werden die Fehler in *Systemfehler* und *Sensorfehler* unterteilt. Ein Systemfehler, z.B. ein blockierendes Ventil, beeinflusst direkt die Systemeigenschaften und wird hier als additiver Term modelliert. Ein Sensorfehler bedeutet, dass der Sensor entweder keine oder fehlerhafte Messwerte liefert. Es wird angenommen, dass zur gleichen Zeit immer nur ein Fehler auftritt, d.h. entweder ein System- oder ein Sensorfehler.

Fehlerdiagnose Aufgabe der Fehlerdiagnose ist, mit Hilfe der erhaltenen Messinformationen auftretende Fehler zu erkennen. Hierbei wird zwischen der *Fehlerdetektion* und der *Fehlerisolation* unterschieden. Bei der Detektion wird erfasst, ob ein Fehler aufgetreten ist. Bei der Isolation wird bei Auftritt eines Fehlers bestimmt, um welchen Fehler es sich handelt bzw. die fehlerhafte Komponente des Systems ermittelt.

Residuum und Residuengenerator Ein Residuum ist eine Größe, die die Abweichung des Verhaltens der realen Strecke vom Modell anzeigt. Der sogenannte *Residuengenerator* errechnet diese aus den gemessenen Ein- und Ausgangsgrößen. In der hier durchgeführten Untersuchung wird als Residuengenerator eine sogenannte Beobachterbank benutzt (siehe Abbildung 4.23). Die Residuen entsprechen den Differenzen zwischen den gemessenen Größen x_i und den durch den Beobachter ermittelten Größen \hat{x}_i . Diese werden in binäre

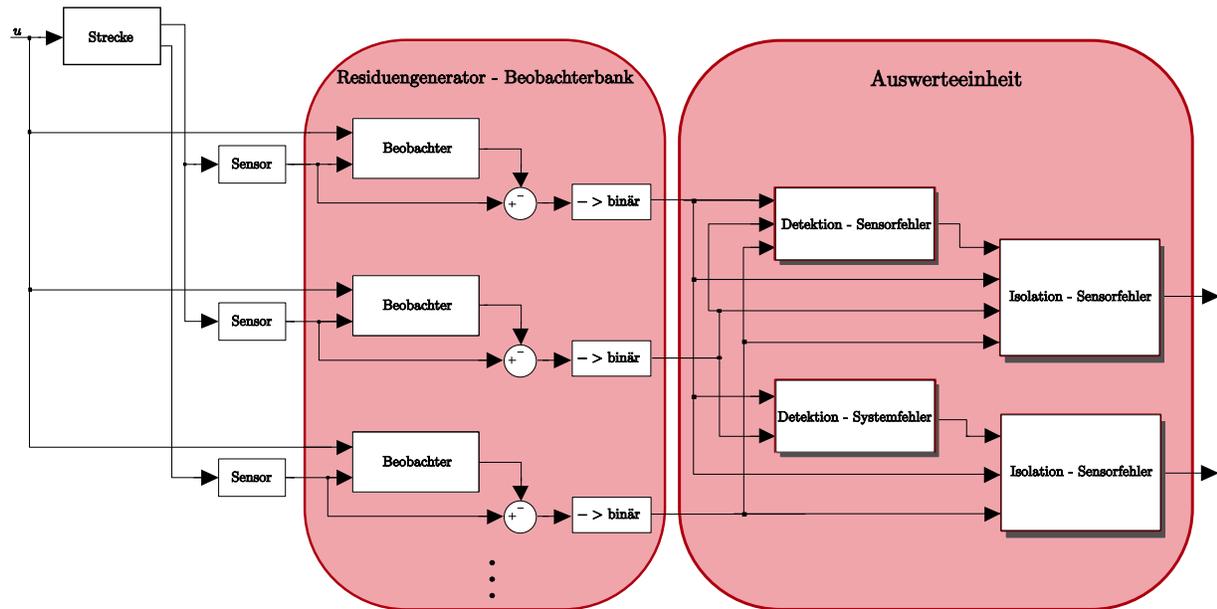


Abbildung 4.23: Strecke mit Diagnoseeinheit

Größen umgewandelt, indem überprüft wird, ob vorgegebene Schranken \tilde{r}_i überschritten werden:

$$r_i = \begin{cases} 0 & \text{für } |x_i - \hat{x}_i| \leq \tilde{r}_i \\ 1 & \text{für } |x_i - \hat{x}_i| > \tilde{r}_i . \end{cases}$$

Diagnoseeinheit Die Diagnoseeinheit besteht aus einem Residuengenerator und Auswerteeinheiten, die anhand der Residuen bestimmen, ob und welcher Fehler aufgetreten ist (siehe Abbildung 4.23).

Die Aufgabe besteht nun darin, festzustellen welche Zustandsvariablen gemessen werden müssen, um eine Fehlerdetektion bzw. -isolation mit Hilfe einer Beobachterbank zu ermöglichen.

4.3.2.1 Fehlerdiagnose mit Hilfe eines Steuerbarkeitsmaß

An dem aus [53] entnommenen System soll gezeigt werden, wie die Maßzahlen bei der Wahl der Sensorplatzierung zur Fehlerdiagnose genutzt werden können. Gegeben sei das System

$$\frac{dx}{dt} = \begin{bmatrix} -1 & -1 & 0 & -1 & -3 \\ 3 & -2 & -4 & -2 & 1 \\ -1 & 0 & -3 & 2 & -4 \\ 5 & 5 & -3 & -4 & -2 \\ -4 & 2 & 3 & 4 & -5 \end{bmatrix} \mathbf{x} + \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} u + \begin{bmatrix} f_1 \\ f_2 \\ 0 \\ f_3 \\ 0 \end{bmatrix} .$$

Es wirken drei Fehler f_i ($i = 1, 2, 3$) auf das System, die als zusätzlicher Summand f_i in den Gleichungen 1, 2 und 4 modelliert werden. Die Idee besteht nun darin, diese Fehler als Eingangsgrößen des Systems zu interpretieren. Hierfür betrachten wir obiges System (für $u = 0$) und erhalten mit der Abkürzung $\mathbf{f} := [f_1, f_2, f_3]^T$ die Systemdarstellung

$$\frac{dx}{dt} = \mathbf{A}\mathbf{x} + \mathbf{B}_f\mathbf{f}$$

mit

$$\mathbf{B}_f = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} =: [\mathbf{b}_1 \quad \mathbf{b}_2 \quad \mathbf{b}_3].$$

Durch Betrachtung des Steuerbarkeitsmaßes m nach BENNINGER (siehe Abschnitt 2.3.2) für jede Spalte von \mathbf{B}_f (d.h. für dieses Beispiel für die Systeme $(\mathbf{A}, \mathbf{b}_1)$, $(\mathbf{A}, \mathbf{b}_2)$ und $(\mathbf{A}, \mathbf{b}_3)$) wird ersichtlich, welche Zustandsvariablen vom jeweiligen Fehler beeinflusst werden. Der Zusammenhang zwischen den möglichen Fehlern und deren Auswirkung auf die Zustandsvariablen ist in nachfolgender Tabelle dargestellt.

Tabelle 4.1: Steuerbarkeitsmaße für die Fehler f_i

	x_1	x_2	x_3	x_4	x_5
$(\mathbf{A}, \mathbf{b}_1) \hat{=} \mathbf{f}_1$	0,3359	0,0226	0,0371	0,6390	0,0291
$(\mathbf{A}, \mathbf{b}_2) \hat{=} \mathbf{f}_2$	0,1807	0,3359	0,8804	0,5867	0,5009
$(\mathbf{A}, \mathbf{b}_3) \hat{=} \mathbf{f}_3$	0,1513	0,0107	0,0169	0,3359	0,0139

Jede Zeile steht für einen Fehler f_i und jede Spalte für eine Zustandsvariable x_i . Die erste Spalte der Tabelle gibt somit an, wie stark die drei Fehler auf die Zustandsvariable x_1 wirken. Da kein Eintrag der Tabelle Null ist, wirkt jeder Fehler auf jede Zustandsvariable. Für eine kompaktere Schreibweise werden die Einträge der Tabelle spaltenweise normiert und in der so genannten Auswirkungsmatrix \mathbf{M} zusammengefasst.

$$\mathbf{M} = \begin{bmatrix} 1 & 0,0673 & 0,0422 & 1 & 0,0581 \\ 0,5379 & 1 & 1 & 0,9182 & 1 \\ 0,4505 & 0,0318 & 0,0192 & 0,5257 & 0,0277 \end{bmatrix}.$$

Anschließend wird eine Schranke angegeben, ab welcher die Wirkung des Fehlers vernachlässigt werden kann. Bei der Wahl der Schranke muss darauf geachtet werden, dass einerseits der Unterschied zwischen den Werten die vernachlässigt und denen die berücksichtigt werden möglichst groß ist. Andererseits muss versucht werden möglichst viele linear unabhängige Spalten zu erzeugen, da diese für die Fehlerisolation benötigt werden. Für dieses Beispiel wurde die Schranke gleich 0,5 gewählt. Damit erhält man die binäre Auswirkungsmatrix

$$\mathbf{M}_b = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}. \quad (4.4)$$

Mit Hilfe dieser Matrix kann nun die Menge der zu messenden Zustandsvariablen für die Fehlerdetektion bzw. -isolation ermittelt werden. Für die Detektion eines Fehlers muss eine Zustandsvariable gemessen werden, auf welche jeder Fehler eine Auswirkung hat. Da die Matrix \mathbf{M}_b zeigt, dass Fehler 3 nur auf x_4 wirkt, ist dies die einzig mögliche Wahl. Weiters kann noch in Betracht gezogen werden, dass die Sensoren fehlerhafte Messwerte liefern, wodurch die Residuen ebenfalls beeinflusst werden. Sensorfehler können durch redundante Messungen detektiert werden. D.h., eine Zustandsvariable wird mit zwei verschiedenen Sensoren gemessen (siehe Abbildung 4.23). Da für dieses Beispiel zur Systemfehlerdetektion x_4 gemessen werden

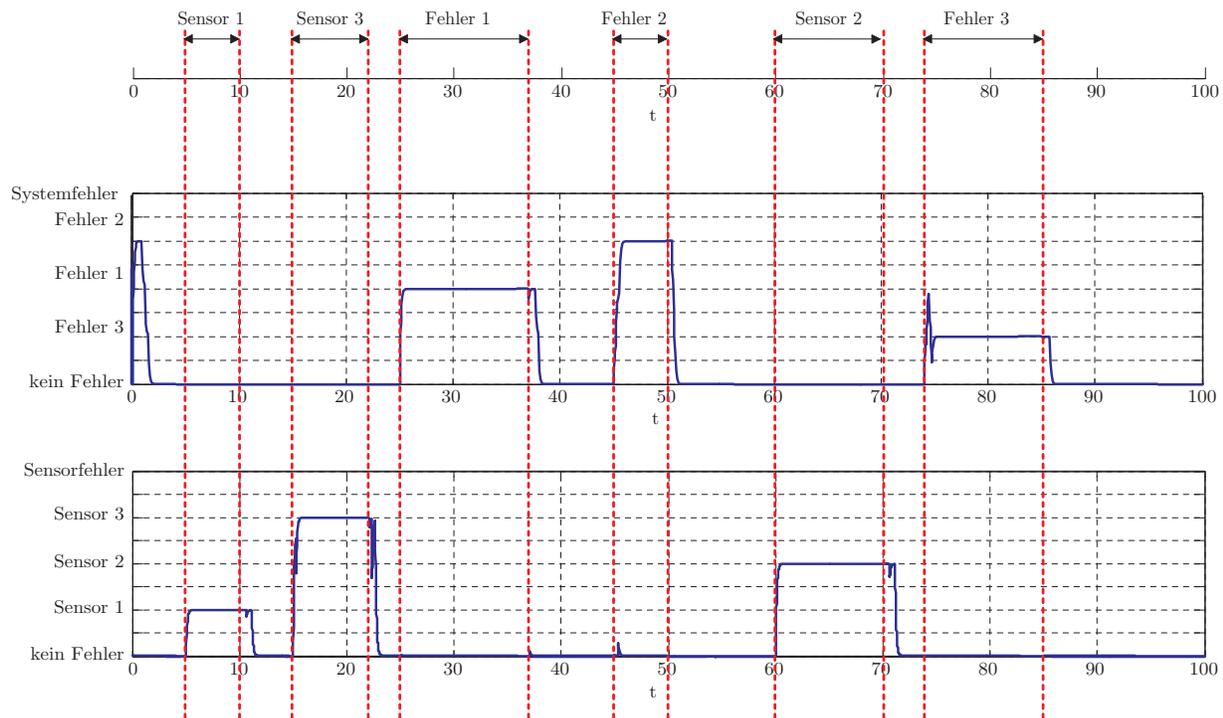


Abbildung 4.24: Ergebnisse der Fehlerdiagnose

muss, ist für die Detektion von System- und Sensorfehler die Menge $\{x_4, x_4\}$ jene, welche am wenigsten Sensoren benötigt.

Wenn bereits sichergestellt wurde, dass ein Systemfehler aufgetreten ist, müssen für die Fehlerisolierung aus der Menge der Zustandsvariablen Residuen ermittelbar sein, die sich bei jedem Fehler eindeutig voneinander unterscheiden. D.h., die zu den Zustandsvariablen gehörenden Spalten der Matrix M_b müssen voneinander unterscheidbar sein. Für das hier betrachtete Beispiel ergeben sich die Mengen $\{x_1, x_2\}$, $\{x_1, x_3\}$ oder $\{x_1, x_5\}$. Die Auswirkungsmatrix M zeigt, dass sich Fehler 2 an der Zustandsvariablen x_3 ein wenig deutlicher von den beiden übrigen Fehlern abhebt als bei x_2 und x_5 . Deshalb fällt die Wahl bei der nachfolgenden Simulation auf die Menge $\{x_1, x_3\}$. Durch die zusätzlich anzubringenden Sensoren können Sensorfehler detektiert und isoliert werden. Für eine rechnerunterstützte Ermittlung möglicher Mengen von Zustandsvariablen siehe [53].

Abbildung 4.24 zeigt die Simulationsergebnisse für die Fehlerdiagnose unter Berücksichtigung von Sensorfehlern. Die Residuen werden aus der zuvor ermittelten Menge $\{x_1, x_3, x_4, x_4\}$ erhalten. Es werden zu unterschiedlichen Zeitpunkten unterschiedlich lang andauernde System- als auch Sensorfehler simuliert. Zu Beginn der Simulation sind die Residuen auch im fehlerfreien Fall ungleich Null, da der Beobachterfehler erst abklingen muss. Die Fehler werden durch die ermittelte Menge von Zustandsvariablen zufrieden stellend isoliert. Auch hier muss eine Einschwingphase abgewartet werden. Ebenfalls muss nach dem Wegfallen des Fehlers eine gewisse Zeitspanne eingehalten werden. Erst danach ist wieder eine zuverlässige Fehlerdiagnose möglich.

In der Literatur werden andere Methoden zur Ermittlung geeigneter Mengen von zu messenden Zustandsvariablen vorgeschlagen (siehe z.B. [26]). Gegenüber diesen Methoden bietet die Fehlerdiagnose mittels Steuerbarkeitsmaßen folgende Vorteile:

- Man kommt schnell und einfach zu einer Lösung.
- Die Maße bieten einen tieferen Einblick in das Systemverhalten, d.h., es kann festgestellt werden, welcher Fehler welches Teilsystem beeinflusst.
- In obigem Beispiel führen die drei Messungen $\{x_1, x_2\}$, $\{x_1, x_3\}$ oder $\{x_1, x_5\}$ zu Fehlerisolierbarkeit. Die durch die Steuerbarkeitsmaße zusätzlich erhaltene Information, wie stark die Fehler auf die einzelnen Variablen wirken, ermöglicht die Auswahl der bestmöglichen Messung bezüglich der Fehlerempfindlichkeit.

4.3.3 Bestimmung des optimalen Eingangsvektors

In diesem Abschnitt wird als weitere Anwendung für Steuerbarkeitsmaße die Wahl des „optimalen“ Eingangsvektors untersucht. Als optimaler Eingangsvektor \mathbf{b} wird derjenige Vektor bezeichnet, der dazu führt, dass das System (\mathbf{A}, \mathbf{b}) bestmöglich steuerbar ist. Durch die Bestimmung dieses Vektors können beispielsweise Sensoren bzw. Aktoren bestmöglich platziert werden.

Für diese Anwendung eignet sich der Ansatz über die minimale Energie (siehe Abschnitt 2.3) auf Grund der einfachen Definition. Im Folgenden wird beschrieben, wie der optimale Eingangsvektor derart bestimmt werden kann, dass die minimale Energie, die notwendig ist, um einen Anfangszustand in den Ursprung zu überführen minimal wird. Um Beschränkungen im Eingangsvektor \mathbf{b} zu berücksichtigen wird das Optimierungsproblem unter der Nebenbedingung

$$\mathbf{bRb} \leq 1$$

gelöst. Hierbei ist \mathbf{R} eine symmetrische, positiv definite (n, n) -Matrix. Damit lässt sich das Optimierungsproblem folgendermaßen zusammenfassen:

$$\min_{\mathbf{b}} \{ |\det(\mathbf{Q}_S^{-1})| \} \quad \text{unter } \mathbf{bRb} \leq 1 \quad \mathbf{R} > 0 \quad (4.5)$$

Hierbei entspricht \mathbf{Q}_S der GRAMschen Steuerbarkeitsmatrix nach Gleichung (2.10). Für das etwas allgemeiner definierte Maß

$$m_k = \left(\frac{1}{n} \text{spur}(\mathbf{Q}_S)^k \right)^{1/k}$$

untersuchten MÜLLER und WEBER in [37] das Optimierungsproblem

$$\max_{\mathbf{b}} \{ m_k \} \quad \text{unter } \mathbf{bRb} \leq 1 \quad \mathbf{R} > 0.$$

Für $k = 0$ entspricht dieses Optimierungsproblem dem hier betrachteten Problem (4.5). In [37] und in [22] wurde das Problem mit Hilfe einer LAGRANGE Funktion gelöst. Verwendet man

für den hier betrachteten speziellen Fall ($k = 0$) den von MÜLLER und WEBER in [37] gewählten Lösungsansatz und führt zusätzlich eine geschickte Transformation des Systems durch (siehe [22]), so kann die Lösung des Optimierungsproblems wesentlich vereinfacht werden. Hierfür wird die Matrix \mathbf{Q}_S durch die Beziehung

$$\mathbf{Q}_S = \mathbf{S}_u \mathbf{F} \mathbf{S}_u^T \quad (4.6)$$

ausgedrückt. Für die Elemente der Matrix \mathbf{F} gilt

$$f_{i,j} = \int_0^{t_1} \alpha_i(-\tau) \alpha_j(-\tau) d\tau.$$

Die Parameter $\alpha_i(t)$ können aus der Relation

$$\Phi(t) = e^{\mathbf{A}t} = \sum_{i=0}^{n-1} \alpha_i(t) \mathbf{A}^i$$

ermittelt werden. Durch die Beziehung (4.6) erhält man für die Determinante der Inversen GRAM-schen Matrix

$$\det(\mathbf{Q}_S^{-1}) = \frac{1}{\det(\mathbf{S}_u)^2 \det(\mathbf{F})}.$$

Da die Matrix \mathbf{F} von \mathbf{b} unabhängig ist, kann das Optimierungsproblem (4.5) beschrieben werden durch

$$\max_{\mathbf{b}} \{|\det(\mathbf{S}_u)|\} \quad \text{unter} \quad \mathbf{b}^T \mathbf{R} \mathbf{b} \leq 1.$$

Da gilt

$$|\det(\mathbf{S}_u(\alpha \mathbf{b}))| > |\det(\mathbf{S}_u(\mathbf{b}))| \quad \text{für} \quad \alpha > 1$$

befindet sich die optimale Lösung am Rand der Nebenbedingung $\mathbf{b}^T \mathbf{R} \mathbf{b}$ und die Ungleichung kann durch die Gleichung $\mathbf{b}^T \mathbf{R} \mathbf{b} = 1$ ersetzt werden [22]. Um das Problem weiter zu vereinfachen wird das System (\mathbf{A}, \mathbf{b}) durch die Zustandstransformation $\mathbf{x} = \mathbf{T} \mathbf{z}$ in ein System (\mathbf{J}, \mathbf{f}) in JORDAN-Normalform überführt. Die Spalten der Transformationsmatrix \mathbf{T} entsprechen den Eigen- bzw. Hauptvektoren der Systemmatrix \mathbf{A} . Die Steuerbarkeitsmatrix des ursprünglichen Systems (\mathbf{A}, \mathbf{b}) lässt sich somit ausdrücken durch

$$\mathbf{S}_u = \mathbf{T} \tilde{\mathbf{S}}_u$$

mit $\tilde{\mathbf{S}}_u$ der Steuerbarkeitsmatrix des transformierten Systems (\mathbf{J}, \mathbf{f}) . Durch die Transformation ergibt sich für die Nebenbedingung

$$\mathbf{f}^T \tilde{\mathbf{R}} \mathbf{f} = 1 \quad \text{mit} \quad \tilde{\mathbf{R}} = \mathbf{T}^T \mathbf{R} \mathbf{T}.$$

Da die Transformationsmatrix \mathbf{T} unabhängig vom Eingangsvektor \mathbf{b} ist, ist es für das Optimierungsproblem ausreichend, die Steuerbarkeitsmatrix $\tilde{\mathbf{S}}_u$ zu betrachten. Der Einfachheit halber wird nun vorerst angenommen, dass die Matrix \mathbf{A} n von einander verschiedene Eigenwerte s_i ($i = 1, \dots, n$) besitzt. Damit entspricht die Transformation der Modaltransformation mit $\mathbf{T} = \mathbf{X}_R$ der Matrix der Rechts-Eigenvektoren, und die Matrix \mathbf{J} wird zur Diagonalmatrix

$\mathbf{J} = \text{diag}(s_i)$. Für diesen Fall kann die Steuerbarkeitsmatrix des transformierten Systems an-
geschrieben werden als das Produkt

$$\tilde{\mathbf{S}}_u = \text{diag}(f_i)\mathbf{V} \quad \text{mit} \quad \mathbf{T}^{-1}\mathbf{b} = \mathbf{f} = \begin{bmatrix} f_1 \\ \vdots \\ f_n \end{bmatrix}$$

mit der sogenannten VANDERMONDE-Matrix

$$\mathbf{V} = \begin{bmatrix} 1 & s_1 & s_1^2 & \cdots & s_1^{n-1} \\ 1 & s_2 & s_2^2 & \cdots & s_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & s_n & s_n^2 & \cdots & s_n^{n-1} \end{bmatrix}.$$

Da auch die Matrix \mathbf{V} unabhängig vom Eingangsvektor \mathbf{b} ist, ergibt sich für das Optimierungs-
problem

$$\max_{\mathbf{f}} \left\{ |\det(\tilde{\mathbf{S}}_u)| \right\} = \max_{\mathbf{f}} \left\{ |\det(\text{diag}(f_i))| \right\} = \max_{\mathbf{f}} \left\{ \prod_{i=1}^n |f_i| \right\} \quad \text{unter} \quad \mathbf{f}^T \tilde{\mathbf{R}} \mathbf{f} = 1.$$

Zur Lösung des Optimierungsproblems wird nun die LAGRANGE Funktion

$$L := \prod_{i=1}^n |f_i| - \lambda(\mathbf{f}^T \tilde{\mathbf{R}} \mathbf{f} - 1)$$

mit dem LAGRANGEN Multiplikator λ eingeführt. Für den transformierten Eingangsvektor \mathbf{f} ,
der das Problem maximiert, gilt die notwendige Bedingung

$$\frac{\partial L}{\partial \mathbf{f}} \stackrel{!}{=} 0 = \frac{\partial}{\partial \mathbf{f}} \left(\prod_{i=1}^n |f_i| \right) - \lambda 2 \tilde{\mathbf{R}} \mathbf{f}. \quad (4.7)$$

Multipliziert man diese Gleichung von links mit \mathbf{f}^T erhält man einen Ausdruck für λ :

$$0 = \mathbf{f}^T \frac{\partial}{\partial \mathbf{f}} \left(\prod_{i=1}^n |f_i| \right) - \lambda \underbrace{\mathbf{f}^T \tilde{\mathbf{R}} \mathbf{f}}_{=1} \quad \leadsto \quad \lambda = \frac{1}{2} \mathbf{f}^T \frac{\partial}{\partial \mathbf{f}} \left(\prod_{i=1}^n |f_i| \right).$$

Setzt man dieses Ergebnis in die notwendige Bedingung (4.7) ein, erhält man den Ausdruck

$$\frac{\partial}{\partial \mathbf{f}} \left(\prod_{i=1}^n f_i \right) - \mathbf{f}^T \frac{\partial}{\partial \mathbf{f}} \left(\prod_{i=1}^n f_i \right) \tilde{\mathbf{R}} \mathbf{f} = 0.$$

Da es sich bei dem Term $\prod_{i=1}^n f_i$ um eine homogene Funktion vom Grad n handelt, gilt nach
dem Satz von EULER über homogene Funktionen

$$n \prod_{i=1}^n f_i = \mathbf{f}^T \frac{\partial}{\partial \mathbf{f}} \left(\prod_{i=1}^n f_i \right).$$

Weiters gilt

$$\frac{\partial}{\partial \mathbf{f}} \left(\prod_{i=1}^n f_i \right) = \prod_{i=1}^n f_i \begin{bmatrix} \frac{1}{f_1} \\ \vdots \\ \frac{1}{f_n} \end{bmatrix}.$$

Aus diesen Beziehungen erhält man nach einigen Umformungen schlussendlich das nichtlineare Gleichungssystem

$$\mathbf{R}\mathbf{b} - \frac{1}{n} \mathbf{T}^{-T} \begin{bmatrix} \frac{1}{f_1} \\ \vdots \\ \frac{1}{f_n} \end{bmatrix} = 0 \quad \text{mit} \quad \mathbf{T}\mathbf{f} = \mathbf{b}.$$

Daraus kann, i. A. numerisch zuverlässig, der optimale Eingangsvektor \mathbf{b} ermittelt werden.

Im Folgenden wird das Optimierungsproblem für den Fall von mehrfachen Eigenwerten der Matrix \mathbf{A} betrachtet. Hierbei sei m_i die Vielfachheit des Eigenwertes s_i , wobei gelte $i = 1, \dots, r$ mit r der Anzahl der von einander verschiedenen Eigenwerte von \mathbf{A} . Für eine im Weiteren einfache Beschreibung des Problems wird der transformierte Eingangsvektor wie folgt zerlegt

$$\mathbf{f}^T = [f_{1,1}, \dots, f_{1,m_1}, f_{2,1}, \dots, f_{2,m_2}, \dots, f_{r,1}, \dots, f_{r,m_r}].$$

Für diesen Fall ist eine ähnliche Faktorisierung der Steuerbarkeitsmatrix $\tilde{\mathbf{S}}_u$ des transformierten System möglich

$$\tilde{\mathbf{S}}_u = \text{diag}(\mathbf{F}_i) \mathbf{V} \quad \text{für} \quad i = 1, \dots, r.$$

Die Struktur der modifizierten Matrix \mathbf{V} kann beispielsweise in [22] nachgelesen werden. Sie ist für das Optimierungsproblem irrelevant, da \mathbf{V} von \mathbf{f} unabhängig ist. Für die Matrizen \mathbf{F}_i gilt

$$\mathbf{F}_i = \begin{bmatrix} f_{i,1} & f_{i,2} & \dots & f_{i,m_i} \\ f_{i,2} & & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ f_{i,m_i} & 0 & \dots & 0 \end{bmatrix}$$

mit der Determinante

$$\det(\mathbf{F}_i) = -f_{i,m_i}^{m_i}.$$

Für das Optimierungsproblem gilt somit

$$\max_{\mathbf{f}} \left\{ |\det(\tilde{\mathbf{S}}_u)| \right\} = \max_{\mathbf{f}} \left\{ |\det(\mathbf{F}_1) \dots \det(\mathbf{F}_r)| \right\} = \max_{\mathbf{f}} \left\{ |(-1)^r \prod_{i=1}^r f_{i,m_i}^{m_i}| \right\} \quad \text{unter} \quad \mathbf{f} \tilde{\mathbf{R}} \mathbf{f} = 1.$$

Analog zur oben beschriebenen Vorgangsweise erhält man die notwendige Bedingung

$$\frac{\partial}{\partial \mathbf{f}} \left((-1)^r \prod_{i=1}^r f_{i,m_i}^{m_i} \right) - n \tilde{\mathbf{R}} \mathbf{f} (-1)^r \prod_{i=1}^r f_{i,m_i}^{m_i} = 0.$$

Für den Lösungsvektor dieses Problems gilt $f_{i,j} = 0$ mit $i = 1, \dots, r$ und $j = 1, \dots, m_i - 1$, da die Ableitung nach diesem Element zu Null wird. Somit verbleiben r nichtlineare Gleichungen für die übrigen Elemente des Vektors \mathbf{f} , die numerisch bestimmt werden.

Beispiel 4.1: Wir betrachten zwei Systeme, für die eine analytische Lösung des Optimierungsproblems existiert.

System in Diagonalform: Gegeben sei das System

$$\frac{dx}{dt} = \text{diag}(s_i)x + \mathbf{b}u \quad \text{mit } i = 1, \dots, n \quad \text{und } \mathbf{b}^T = [b_1, \dots, b_n].$$

Für die Lösung des Optimierungsproblem erhält man mit der Gewichtungsmatrix $\mathbf{R} = \text{diag}(r_i)$ die n notwendigen Bedingungen

$$r_i b_i = \frac{1}{nb_i},$$

woraus folgt

$$b_i = \frac{1}{\sqrt{nr_i}} \quad i = 1, \dots, n.$$

Der optimale Eingangsvektor ist unabhängig von der Systemmatrix \mathbf{A} und die Elemente dieses Vektors sind für den Fall, dass die einzelnen Zustandvariablen gleich gewichtet sind, alle gleich groß.

System in JORDAN-Form: Gegeben sei

$$\frac{dx}{dt} = \begin{bmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ & & \lambda & 1 \\ 0 & & & \lambda \end{bmatrix} \mathbf{x} + \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} u \quad \text{und } \mathbf{R} = \text{diag}(r_1, \dots, r_n).$$

Aus den notwendigen Bedingungen

$$\frac{\partial}{\partial \mathbf{b}} (-b_n^n) + nb_n^n \mathbf{R} \mathbf{b} = 0$$

ergibt sich für den optimalen Eingangsvektor

$$\mathbf{b}^T = \left[0, \dots, 0, \frac{1}{\sqrt{r_n}} \right].$$

Liegt das System in JORDAN-Form vor, so ist für den optimalen Eingangsvektor nur der n -te Eintrag entscheidend.

4.3.3.1 Bemerkungen zum Distanzmaß

Aufgabenstellungen dieser Art können auch mit Hilfe des Distanzmaßes gelöst werden. Allerdings muss hier über das i.A. bereits aufwendige Optimierungsproblem zur Berechnung des Distanzmaßes ein weiteres Optimierungsproblem „gestülpt“ werden:

$$\max_{\mathbf{b}} \left\{ \inf_{s \in \mathbb{C}} \{ \sigma_{\min}([s\mathbf{E} - \mathbf{A}, \mathbf{B}]) \} \right\}.$$

Analog können auf diesen Weg auch Parameter der Systemmatrix \mathbf{A} bezüglich des Steuerbarkeitsmaßes optimiert werden, wie nachfolgendes Beispiel aus [6] zeigt.

Beispiel 4.2: Gegeben sei das System

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 2 & 3 \\ -1 & 1 & 4 & 5 \\ 0 & x_1 & 1 & 2 \\ x_2 & 0 & -2 & 1 \end{bmatrix} \quad \text{und} \quad \mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}.$$

Hierbei sind x_1 und x_2 frei wählbare, reelle Konstanten. Für die Wahl $x_1 = x_2 = 1$ ergibt sich das Distanzmaß $\mu_C = 0,19$. Durch Optimierung über x_1 und x_2 kann ein lokales Minimum für $x_1 = 1,98$ und $x_2 = -1,87$ mit dem Wert $\mu_C = 0,49$ ermittelt werden. In Abbildung 4.25 sind die Pseudospektren für diese beiden Fälle dargestellt.

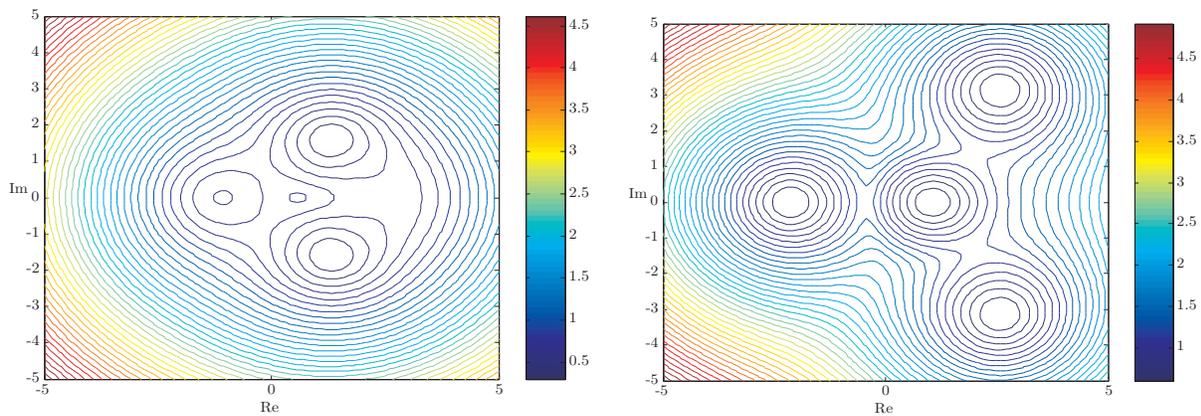


Abbildung 4.25: Pseudospektrum für $x_1 = x_2 = 1$ (links) und für $x_1 = 1,98$ und $x_2 = -1,87$ (rechts)

Kapitel 5

Anwendung der Maße auf zeitdiskrete Systeme

In der Praxis kommt den zeitdiskreten, linearen und zeitinvarianten Systemen beim Entwurf von Abtastregelkreisen eine große Bedeutung zu. Die Struktur der Kriterien für die Steuer- bzw. Erreichbarkeit von zeitdiskreten Systemen stimmen weitgehend mit denen für zeitkontinuierliche Systeme überein, weshalb die eingeführten Maße auch auf diese Systeme angewendet werden können. Dadurch ergibt sich i.A. die gleiche Auswertung der Maße für diskrete Systeme. In diesem Kapitel wird geklärt, inwiefern die für zeitkontinuierliche Systeme entwickelten Maße auch für zeitdiskrete Systeme geeignet sind.

Als mögliche Anwendung der Maße in diesem Bereich wird die Wahl der Abtastzeit für zeitdiskretisierte Systeme untersucht. Abgesehen von physikalischen Einschränkungen gibt es für die Wahl der Abtastzeit eine von KALMAN formulierte Einschränkung (vgl. Abschnitt 5.2.2). Dieses Kapitel beschäftigt sich mit der Frage, ob es nicht nur *nicht erlaubte* Werte für T_d gibt, sondern auch eine „gute“ bzw. „schlechte“ Wahl für die Abtastzeit, die die Steuerbarkeit des zeitdiskreten Systems beeinflusst. Für die durchgeführten Untersuchungen werden nur ausgewählte Maße, angeführt in Abschnitt 5.1, betrachtet, welche sich durch die Analyse der vorangegangenen Kapitel als zuverlässig und effizient berechenbar erwiesen haben.

Wie in den folgenden Abschnitten gezeigt wird, liefern die in dieser Arbeit betrachteten Maße sehr unterschiedliche Resultate. Daher werden die ermittelten Ergebnisse in Abschnitt 5.3.1 mit Hilfe eines diskreten Zustandsreglerentwurfs genauer untersucht. Da sowohl mit Hilfe der Verstärkung des Zustandsreglers als auch mittels der benötigten Stelleistung Aussagen über die „Güte“ der Steuerbarkeit getroffen werden können, können diese mit den aus den Maßzahlen erhaltenen Erkenntnissen verglichen werden (siehe Abschnitt 5.3.1). Abschließend wird die Anwendbarkeit der Maßzahlen auf diskrete Systeme an den Beispielen „Wagen mit Stab“, „Hubmagnet“ und „Doppelpendel“ überprüft und die Ergebnisse zusammengefasst.

5.1 Diskretisierungsmethode und Maßzahlen

Diskreter Simulator: Werden für das zeitdiskrete System

$$\xi_{i+1} = \mathbf{A}_d \xi_i + \mathbf{B}_d \mathbf{u}_i \quad \eta_i = \mathbf{C}_d \xi_i$$

die Matrizen

$$\mathbf{A}_d = e^{\mathbf{A}T_d} \quad \text{und} \quad \mathbf{B}_d = \int_0^{T_d} e^{\mathbf{A}\tau} \mathbf{B} d\tau \quad (5.1)$$

gewählt, dann gilt

$$\boldsymbol{\xi}_i = \mathbf{x}(iT_d)$$

für eine stückweise konstante Eingangsfunktion

$$u(\tau) = u(iT_d) \quad \text{für} \quad iT_d \leq \tau < (i+1)T_d \quad \text{mit} \quad i \geq 0.$$

Durch die weitere Wahl

$$\mathbf{C}_d = \mathbf{C} \quad (5.2)$$

erreicht man, dass die Beziehung

$$\boldsymbol{\eta}_i = \mathbf{y}(iT_d) \quad \text{für} \quad i \geq 0 \quad (5.3)$$

gilt. Hierbei ist T_d die sogenannte Abtast- oder Diskretisierungszeit mit der natürlichen Eigenschaft $T_d > 0$. Ein System mit den Eigenschaften (5.1) und (5.3) nennt man *diskreten Simulator*.

Maßzahlen: Für die durchzuführenden Untersuchungen wird als modales Steuerbarkeitsmaß das Maß κ nach LITZ

$$\kappa_i = \frac{\mathbf{x}_{Li}^H \mathbf{B}_d \mathbf{B}_d^T \mathbf{x}_{Li}}{\mathbf{x}_{Li}^H \mathbf{x}_{Li}}$$

herangezogen.

Für die Beurteilung der Steuerbarkeit des Gesamtsystems dient das Distanzmaß

$$\mu_C(\mathbf{A}_d, \mathbf{B}_d) = \inf_{s \in \mathbb{C}} \sigma_{\min}([s\mathbf{E} - \mathbf{A}_d, \mathbf{B}_d]). \quad (5.4)$$

Für dessen Berechnung wird in den folgenden Abschnitten der Trisection-Algorithmus benutzt, da sich dieser als sehr zuverlässig erwiesen hat.

5.2 Maße in Abhängigkeit der Abtastzeit

Im Folgenden werden die ausgewählten Maße sofern möglich *analytisch* berechnet. Somit erhält man für jede Maßzahl eine Funktion der unabhängigen Variablen T_d , deren Eigenschaften in den folgenden Abschnitten untersucht werden. Durch Betrachtung von Systemen niedriger Ordnung soll Einsicht in das Steuerbarkeitsverhalten von Abtastsystemen gewonnen werden, die es ermöglicht allgemeine Aussagen zu treffen.

Beginnend mit Systemen 1. Ordnung werden die Eigenschaften der Maßfunktionen für den diskreten Simulator diskutiert. In Abschnitt 5.2.2 erfolgt eine analoge Untersuchung für Systeme 2. Ordnung. Die Ergebnisse werden in Abschnitt 5.2.3 soweit möglich auf Systeme höherer Ordnung übertragen.

5.2.1 Systeme 1. Ordnung

Für Systeme 1. Ordnung kann die Abhängigkeit der Maße von der Abtastzeit T_d noch einfach analytisch ermittelt werden. Dadurch ist es möglich, das Verhalten (Monotonie, Grenzwerte, Unstetigkeitsstellen, ...) dieser Funktionen zu untersuchen. Interessant ist die analytische Lokalisierung von Extrema der Funktion, wodurch die optimale Abtastzeit T_d^* ermittelt werden kann, für die das zeitdiskrete System bestmöglich steuerbar ist.

Das Steuerbarkeitsmaß κ nach LITZ und das Distanzmaß μ sind immer positiv, somit ist die optimale Abtastzeit T_d^* genau jene für die $\kappa(T_d)$ bzw. $\mu(T_d)$ ein globales Maximum besitzt.

Aus dem zeitkontinuierlichen, steuerbaren System 1. Ordnung (mit $b \neq 0$)

$$\frac{dx}{dt} = ax + bu \quad y = cx$$

ergibt sich für den diskreten Simulator nach Gleichung (5.1) und (5.2)

$$\xi_{i+1} = e^{aT_d}\xi_i + \frac{b}{a}(e^{aT_d} - 1)u_i =: a_d\xi_i + b_d u_i \quad \eta_i = c\xi_i =: c_d\xi_i. \quad (5.5)$$

Der diskrete Simulator ist genau dann nicht steuerbar, wenn gilt

$$b_d = \frac{b}{a}(e^{aT_d} - 1) = 0.$$

D.h., das System ist für beliebige Abtastzeiten ($T_d > 0$) steuerbar.

5.2.1.1 Steuerbarkeitsmaß κ

Die Berechnung des Steuerbarkeitsmaßes $\kappa(T_d)$ nach LITZ für das diskrete System 1. Ordnung (5.5) ergibt

$$\kappa(T_d) = b_d^2 = \frac{b^2}{a^2} (e^{aT_d} - 1)^2. \quad (5.6)$$

Für die erste Ableitung nach der Abtastzeit T_d gilt

$$\frac{d\kappa}{dT_d} = \frac{2b^2}{a^2} (e^{aT_d} - 1) e^{aT_d} > 0,$$

woraus folgt, dass $\kappa(T_d)$ eine monoton steigende Funktion ist.

Je nach Vorzeichen von a hat die Funktion $\kappa(T_d)$ zwei unterschiedliche Verhalten, dargestellt in Abbildung 5.1. Für $a > 0$ strebt die Funktion $\kappa(T_d)$ für $T_d \rightarrow \infty$ gegen Unendlich und für $a < 0$ ergibt sich der Grenzwert

$$\lim_{T_d \rightarrow \infty} \kappa(T_d) = \frac{b^2}{a^2}.$$

Für $T_d \rightarrow 0$ strebt die Funktion unabhängig vom Vorzeichen von a gegen Null.

Aus diesen Ergebnissen schließt man, dass je größer T_d wird, desto größer wird auch κ , und desto besser steuerbar wird das diskrete System.

5.2.1.2 Distanzmaß μ

Für das zeitdiskrete System 1. Ordnung (5.5) kann das Distanzmaß μ nach Gleichung (5.4) analytisch berechnet werden. Für $s^* = a_d$ wird Gleichung (5.4) zu einem Minimum mit dem Wert

$$\mu(T_d) = \inf_{s \in \mathbb{C}} \sigma_{\min}([s^* - a_d, b_d]) = |b_d| = \left| \frac{b}{a} (e^{aT_d} - 1) \right|.$$

Da $\mu^2 = \kappa$ gilt, weisen die beiden Maße das gleiche qualitative Verhalten auf.

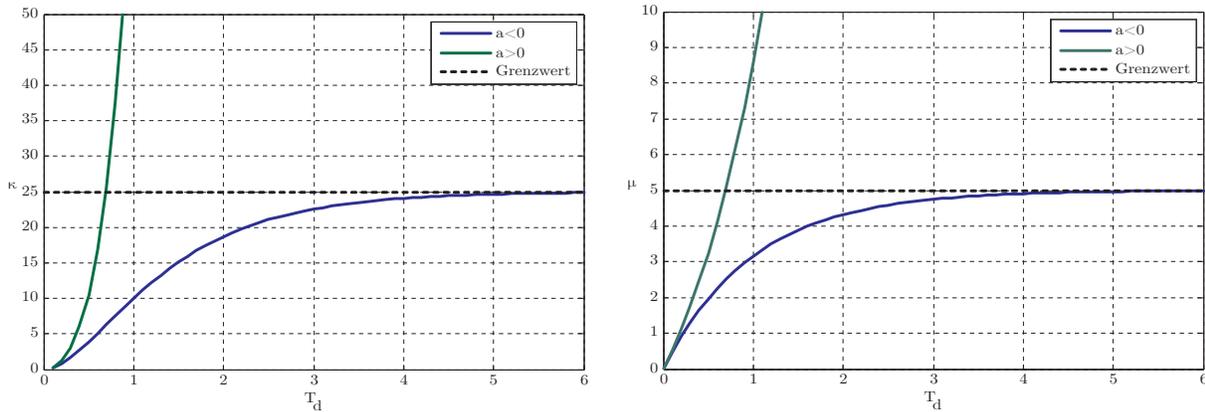


Abbildung 5.1: Steuerbarkeitsmaß $\kappa(T_d)$ (links) und Distanzmaß $\mu(T_d)$ (rechts) für den diskreten Simulator 1. Ordnung

5.2.1.3 Diskussion der Ergebnisse

Für kleiner werdende Abtastzeiten $T_d \rightarrow 0$ gleicht das zeitdiskrete System immer mehr dem kontinuierlichen System. Somit würde man erwarten, dass auch die Maßzahl dem Steuerbarkeitsmaß des kontinuierlichen Systems ($\mu = |b|$ und $\kappa = b^2$) zustrebt. Beide Maße streben jedoch für $T_d \rightarrow 0$ gegen Null. Das bedeutet, dass die Maße für kleine Abtastzeiten eine schlechte Steuerbarkeit anzeigen. Der Grund dafür liegt an dem Wert b_d , welcher mit abnehmender Abtastzeit gegen Null strebt. Dementsprechend wird eine große Eingangsgröße u benötigt um den gewünschten Einfluss auf das System zu bewirken. Dies kann als schlechte Steuerbarkeit des diskreten Systems gedeutet werden.

Zusammenfassung (Diskreter Simulator für Systeme 1. Ordnung):

1. Beide Maße κ und μ streben nach Null für $T_d \rightarrow 0$ und zeigen für kleine Werte der Abtastzeit T_d schlechte Steuerbarkeit an.
2. Für steigende Werte der Abtastzeit T_d nehmen beide Maße betragsmäßig höhere Werte an. D.h., die Güte der Steuerbarkeit nimmt mit größer werdender Abtastzeit zu.
3. Für $a < 0$ kann für die beiden Funktionen $\kappa(T_d)$ und $\mu(T_d)$ eine endliche obere Schranke angegeben werden. Für $a > 0$ existiert keine endliche Schranke.

5.2.2 Systeme 2. Ordnung

Für Systeme der Ordnung $n \geq 2$ muss bei der Wahl der Abtastzeit der Satz von KALMAN beachtet werden:

Theorem 12 (Satz von KALMAN): Ein aus einem steuerbaren und zeitkontinuierlichen System (1.1) entstandenes zeitdiskretes System

$$\xi_{i+1} = e^{\mathbf{A}T_d} \xi_i + \left(\int_0^{T_d} e^{\mathbf{A}\tau} \mathbf{B} d\tau \right) \mathbf{u}_i \quad \eta_i = \mathbf{C} \xi_i$$

ist steuerbar und beobachtbar, wenn gilt: für alle Eigenwerte

$$\lambda_\mu = \sigma + j\omega_\mu \quad \text{und} \quad \lambda_\nu = \sigma + j\omega_\nu$$

der Matrix \mathbf{A} mit gleichem Realteil erfüllt die Abtastzeit T_d die Ungleichung:

$$(\omega_\mu - \omega_\nu)T_d \neq \pm k2\pi \quad \text{mit } k = 1, 2, \dots \quad (5.7)$$

Für einen skalaren Eingang u ($m = 1$) ist obige Bedingung notwendig und hinreichend. Handelt es sich um eine vektorielle Eingangsgröße \mathbf{u} ($m > 1$), ist obiger Satz nur hinreichend [43].

Nach diesem Satz gibt es für den diskreten Simulator Abtastzeiten T_d , für welche das System die Steuer- und Beobachtbarkeit verliert. Diese Abtastzeit, im Weiteren mit $T_{d,K}$ bezeichnet, lässt sich durch Umformen von Gleichung (5.7) berechnen:

$$T_{d,K} = \frac{\pi}{\omega}. \quad (5.8)$$

Der diskrete Simulator ist für Vielfache dieser Abtastzeit $T_{d,K}$ nicht steuerbar.

Ein Steuerbarkeitsmaß ist für den diskreten Simulator nur dann sinnvoll, wenn es mit dem Satz von KALMAN konsistent ist. Es muss für Vielfache der Abtastzeit $T_{d,K}$ zu Null werden und somit den Steuerbarkeitsverlust korrekt wiedergeben.

Im Folgenden wird der Einfluss der Abtastzeit T_d auf die Steuerbarkeit des aus dem steuerbaren und zeitkontinuierlichen, in Modalform vorliegenden Systems

$$\mathbf{A} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \quad (5.9)$$

entstandenen diskreten Simulator untersucht. Für den diskreten Simulator ergeben sich nach den Gleichungen (5.1) und (5.2) die Systemmatrizen

$$\mathbf{A}_d = \begin{bmatrix} e^{\lambda_1 T_d} & 0 \\ 0 & e^{\lambda_2 T_d} \end{bmatrix} \quad \text{und} \quad \mathbf{b}_d = \begin{bmatrix} \frac{b_1}{\lambda_1} (e^{\lambda_1 T_d} - 1) \\ \frac{b_2}{\lambda_2} (e^{\lambda_2 T_d} - 1) \end{bmatrix}. \quad (5.10)$$

Berechnet man die Steuerbarkeitsmatrix

$$\mathbf{S}_u = [\mathbf{b}_d \quad \mathbf{A}_d \mathbf{b}_d]$$

nach KALMAN, so ergibt sich

$$\mathbf{S}_u = \begin{bmatrix} \frac{b_1}{\lambda_1} (e^{\lambda_1 T_d} - 1) & \frac{b_1}{\lambda_1} (e^{\lambda_1 T_d} - 1) e^{\lambda_1 T_d} \\ \frac{b_2}{\lambda_2} (e^{\lambda_2 T_d} - 1) & \frac{b_2}{\lambda_2} (e^{\lambda_2 T_d} - 1) e^{\lambda_2 T_d} \end{bmatrix}.$$

Für die Determinante dieser Matrix erhält man

$$\det(\mathbf{S}_u) = b_1 b_2 \left[\frac{(e^{\lambda_1 T_d} - 1)}{\lambda_1} \frac{(e^{\lambda_2 T_d} - 1)}{\lambda_2} (e^{\lambda_2 T_d} - e^{\lambda_1 T_d}) \right].$$

Sie ist offensichtlich Null für

$$e^{\lambda_1 T_d} = 1 \quad \text{bzw.} \quad e^{\lambda_2 T_d} = 1$$

oder

$$e^{\lambda_1 T_d} = e^{\lambda_2 T_d}.$$

Letztere Bedingung kann, da $\lambda_1 \neq \lambda_2$ gilt, nur bei komplexen Eigenwerten auftreten. Für $\lambda_{1,2} = \sigma \pm j\omega$ erhält man

$$e^{\lambda_1 T_d} - e^{\lambda_2 T_d} = e^{\sigma T_d} (e^{-j\omega T_d} - e^{j\omega T_d}) = -e^{\sigma T_d} 2j \sin(\omega T_d).$$

Damit ist leicht erkennbar, dass für $T_d = kT_{d,K}$ mit $k = 1, 2, \dots$ obige Determinante zu Null wird.

Da sich die Abhängigkeit der Maße von T_d für $\lambda_1, \lambda_2 \in \mathbb{R}$ und $\lambda_1, \lambda_2 \in \mathbb{C}$ teilweise stark unterscheidet, werden diese Fälle für das jeweilige Maß separat betrachtet.

5.2.2.1 Steuerbarkeitsmaß κ

Da auch das zeitdiskrete System (5.10) in Modalform vorliegt, sind die Einheitsvektoren eine mögliche Wahl für die Eigenvektoren. Damit lässt sich das Modalmaß κ nach LITZ leicht ermitteln und man erhält

$$\kappa_{1,2}(T_d) = \left| \frac{b_{1,2}^2 (-1 + e^{T_d \lambda_{1,2}})^2}{\lambda_{1,2}^2} \right|. \quad (5.11)$$

Wie schon bei Systemen 1. Ordnung gilt auch hier

$$\lim_{T_d \rightarrow 0} \kappa(T_d) = 0.$$

1. Fall: Reelle Eigenwerte Besitzt die Matrix \mathbf{A} zwei reelle Eigenwerte ($\lambda_1 \neq \lambda_2$), so erhält man für $\kappa_1(T_d)$ und $\kappa_2(T_d)$ zwei unterschiedliche, monoton steigende Maßfunktionen für die beiden Eigenwerte. Die Maße entsprechen der Lösung für Systeme 1. Ordnung nach Gleichung (5.6), auf die bereits in Abschnitt 5.2.1.1 eingegangen wurde.

Da beide Maßfunktionen monoton steigend sind, kommt man, wie schon bei Systemen 1. Ordnung, zu dem Ergebnis, dass mit steigender Abtastzeit T_d auch die Steuerbarkeit des Systems besser wird. Somit liegt die optimale Abtastzeit theoretisch im Unendlichen.

Für die Praxis könnten zwei Fragen von Interesse sein, die sich mit dem Steuerbarkeitsmaß κ folgendermaßen beantworten lassen:

- Wie groß ist bei einer vorgegebenen Abtastzeit T_d die Mindest-Güte Γ der Steuerbarkeit?
Für die vorgegebene Abtastzeit T_d errechnet sich die Mindest-Güte zu

$$\Gamma = \min\{\kappa_1(T_d), \kappa_2(T_d)\}.$$

- Wie klein kann die Abtastzeit T_d gewählt werden, wenn eine Mindest-Güte Γ nicht unterschritten werden soll?

Für die vorgegebene Mindest-Güte Γ erhält man

$$T_d = \max \left\{ \frac{1}{\lambda_1} \ln \left(1 + \lambda_1 \sqrt{\frac{\Gamma}{b_1^2}} \right), \frac{1}{\lambda_2} \ln \left(1 + \lambda_2 \sqrt{\frac{\Gamma}{b_2^2}} \right) \right\}. \quad (5.12)$$

Diese Abtastzeit darf nicht unterschritten werden, wenn das System die Mindest-Güte Γ besitzen soll.

2. Fall: Konjugiert komplexe Eigenwerte Für das System (5.9) gelte $\lambda_{1,2} = \sigma \pm j\omega$ und $b_1 = \bar{b}_2 =: b$. Daraus folgt: $\kappa_1 = \kappa_2 =: \kappa$. Man erhält eine einzige Funktion $\kappa(T_d)$, die in Abhängigkeit von σ drei unterschiedliche Verhalten aufweist (siehe Abbildung 5.2).

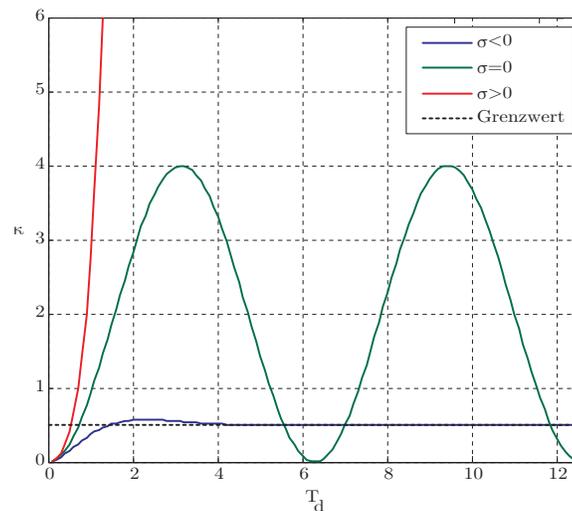


Abbildung 5.2: Steuerbarkeitsmaß κ für konjugiert komplexe Eigenwerte der Matrix **A**

1. Fall $\sigma < 0$: Für $\sigma < 0$ strebt die Funktion $\kappa(T_d)$ dem Grenzwert

$$\lim_{T_d \rightarrow \infty} \kappa(T_d) = \frac{|b|^2}{\sigma^2 + \omega^2}$$

zu. Die Funktion $\kappa(T_d)$ besitzt ein Maximum, das jedoch bisher nicht analytisch ermittelt werden konnte.

2. Fall $\sigma > 0$: Ist $\sigma > 0$, strebt die Funktion $\kappa(T_d)$ für größer werdende Abtastzeiten T_d gegen Unendlich.

In beiden Fällen ($\sigma > 0$ oder $\sigma < 0$) spiegelt die Funktion $\kappa(T_d)$ den Satz von KALMAN (siehe Theorem 12) *nicht wieder*, da $\kappa(T_d = T_{d,K})$ nicht Null ist und somit der Steuerbarkeitsverlust nicht angezeigt wird.

3. Fall $\sigma = 0$: Gilt $\sigma = 0$, vereinfacht sich Gleichung (5.11) zu

$$\kappa(T_d) = \frac{2|b|^2}{\omega^2} (1 - \cos(\omega T_d)). \quad (5.13)$$

Diese Funktion besitzt Nullstellen für $T_d = \frac{2k\pi}{\omega} = 2kT_{d,K}$ ($k = 1, 2, \dots$) und Extrema an den Stellen $T_d = \frac{k\pi}{\omega} = kT_{d,K}$ ($k = 1, 2, \dots$). Dabei handelt es sich bei ungeradzahligem Wert für k um Maxima der Funktion (5.13). Für geradzahliges k erhält man die Minima bzw. die Nullstellen der Funktion $\kappa(T_d)$. Man beachte, dass für $T_d = kT_{d,K}$ ($k = 1, 2, \dots$) nach KALMAN das diskrete System nicht steuerbar ist!

Der Grund für die Inkonsistenz mit dem Satz von KALMAN liegt darin, dass das Steuerbarkeitsmaß κ nach LITZ den Verlust der Steuerbarkeit nur durch das Nullwerden des Eingangsvektors \mathbf{b}_d detektieren kann. Für Vielfache der Abtastzeit $T_{d,K}$ ist der Eingangsvektor mit

$$\mathbf{b}_d(kT_{d,K}) = \begin{bmatrix} \frac{b}{\sigma + j\omega} \\ \frac{b}{\sigma - j\omega} \end{bmatrix} \left((-1)^k e^{k\frac{\sigma\pi}{\omega}} - 1 \right)$$

gegeben. Für $\sigma \neq 0$ kann $|\mathbf{b}_d|$ nicht gleich Null werden. Gilt $\sigma = 0$, so wird der Betrag von \mathbf{b}_d nur für geradzahliges k zu Null. Für $T_d = kT_{d,K}$ ($k = 1, 2, \dots$) liegen allerdings die beiden identischen Eigenwerte der Matrix \mathbf{A}_d

$$e^{(\sigma \pm j\omega)kT_{d,K}} = (-1)^k e^{\frac{\sigma\pi}{\omega}}$$

vor, die zu dem Steuerbarkeitsverlust führen. Da das Steuerbarkeitsmaß κ auf der Modalform basiert, kann das Maß den Steuerbarkeitsverlust nicht anzeigen.

5.2.2.2 Distanzmaß μ

Abschnitt 3.1.2 beinhaltet die Ermittlung einer analytischen Lösung des Distanzmaßes für kontinuierliche Systeme. Die erhaltenen Ergebnisse werden nun auf die diskreten Systeme übertragen. Hierfür werden die Systemdaten des kontinuierlichen Systems (\mathbf{A}, \mathbf{b}) durch die des diskretisierten Systems $(\mathbf{A}_d, \mathbf{b}_d)$ ersetzt.

1. Fall: Reelle Eigenwerte Für ein System 2. Ordnung mit reellen Eigenwerten kann der Wert s^* nicht analytisch berechnet werden, da zur Bestimmung von x^* das Lösen einer Polynomgleichung 4. Ordnung notwendig ist. In Abschnitt 3.1.2.1 wurden für kontinuierliche Systeme Asymptoten zur Abschätzung des Distanzmaßes ermittelt. Führt man für den Abstand zwischen den Eigenwerten die Abkürzung

$$\delta = |e^{\lambda_1 T_d} - e^{\lambda_2 T_d}|$$

ein, ergeben sich für diskrete Systeme die beiden Asymptoten

$$\mu_{a_1} := \frac{1}{2}\delta \quad \text{und} \quad \mu_{a_2} := \min\{b_{d_1}, b_{d_2}\}.$$

Für eine vorgegebene Abtastzeit T_d ergibt sich dadurch analog zu den kontinuierlichen Systemen eine einfache Abschätzung des Distanzmaßes. Eine genaue Analyse der Abhängigkeit der Asymptoten von der Abtastzeit gestaltet sich allerdings schwierig. Durch die Untersuchung der Grenzwerte kann jedoch ein Einblick in das generelle Verhalten dieser Abhängigkeit gewonnen werden: Strebt die Abtastzeit gegen Null, so wird der Abstand δ zwischen den Eigenwerten immer kleiner. Für kleine δ muss die Asymptote μ_{a_1} betrachtet werden. Für diese gilt:

$$\lim_{T_d \rightarrow 0} \mu_{a_1} = 0.$$

Auf analoge Weise wird auch der Grenzwert der Asymptoten für $T_d \rightarrow \infty$ untersucht, woraus man auf die Grenzwerte des Distanzmaßes $\mu(T_d)$ schließen kann:

$$\begin{aligned} \lambda_1 > 0, \lambda_2 > 0 & \quad \lim_{T_d \rightarrow \infty} \mu(T_d) = \infty \\ \lambda_1 < 0, \lambda_2 > 0 & \quad \lim_{T_d \rightarrow \infty} \mu(T_d) = \left| \frac{b_1}{\lambda_1} \right| \\ \lambda_1 > 0, \lambda_2 < 0 & \quad \lim_{T_d \rightarrow \infty} \mu(T_d) = \left| \frac{b_2}{\lambda_2} \right| \\ \lambda_1 < 0, \lambda_2 < 0 & \quad \lim_{T_d \rightarrow \infty} \mu(T_d) = 0. \end{aligned}$$

Diese Ergebnisse konnten durch numerische Berechnung des Distanzmaßes bestätigt werden. Sind beide Eigenwerte positiv, existiert keine endliche optimale Abtastzeit T_d^* . Für die übrigen Fälle sind endliche Werte für die optimale Abtastzeit möglich.

2. Fall: Konjugiert komplexe Eigenwerte Man nimmt an, dass die Matrix \mathbf{A} des kontinuierlichen Systems (5.9) konjugiert komplexe Eigenwerte $\lambda_{1,2} = \sigma \pm j\omega$ besitzt und für den Eingangsvektor $b_1 = \bar{b}_2 := b$ gilt. Damit besitzt das diskrete System konjugiert komplexe Eigenwerte:

$$\Lambda_{1,2} = e^{\sigma T_d} \cos(\omega T_d) \pm j e^{\sigma T_d} \sin(\omega T_d) =: \sigma_d \pm j\omega_d$$

sowie konjugiert komplexe Elemente für den Eingangsvektor

$$b_{d1} = \bar{b}_{d2} := b_d = \frac{b}{\sigma + j\omega} \left(e^{(\sigma + j\omega)T_d} - 1 \right).$$

Für Systeme mit dieser Struktur wurden in Abschnitt 3.1.2.2 zwei mögliche Lösungen des Distanzmaßes analytisch ermittelt. Die Lage eines Minimums wird durch den Realteil der Eigenwerte $\Lambda_{1,2}$

$$s_1^* = \sigma_d = e^{\sigma T_d} \cos(\omega T_d)$$

bestimmt. Der Wert dieses Minimums ergibt sich aus dem Betrag des Imaginärteils zu

$$\mu_1 := \mu|_{s=s_1^*} = e^{\sigma T_d} |\sin(\omega T_d)|.$$

Aus dem konjugiert komplexen Minimum

$$s_{2,3}^* = \sigma_d \pm \frac{1}{2\omega_d} \sqrt{4\omega_d^4 - |b_d|^4}$$

ergibt sich nach Gleichung (3.14) eine weitere mögliche Lösung für das Distanzmaß

$$\mu_2 := \mu|_{s=s_{2,3}^*} = \frac{1}{2} \left| \frac{b_d}{\omega_d} \right| \sqrt{4\omega_d^2 - |b_d|^2}. \quad (5.14)$$

Gültigkeitsbereiche der Lösungen Im Folgenden werden einfache Bedingungen erarbeitet, deren Erfüllung nur noch eine der beiden möglichen Lösungen zulässt. Die Lösung μ_2 ist unzulässig, wenn sie eine komplexe Zahl ist, wodurch μ_1 die einzige gültige Lösung ist. Dies ist genau dann der Fall, wenn die Ungleichung

$$|b_d|^2 > 4\omega_d^2$$

erfüllt ist. Daraus ergibt sich

$$e^{2\sigma T_d} - 2e^{\sigma T_d} \cos \omega T_d + 1 > \frac{4}{K} e^{2\sigma T_d} \sin^2 \omega T_d \quad \text{mit} \quad K := \frac{|b|^2}{\sigma^2 + \omega^2} > 0. \quad (5.15)$$

Obige Ungleichung wird nun mit dem Term $e^{-2\sigma T_d}$ multipliziert und die reelle und nicht negative Variable $x := e^{-\sigma T_d}$ eingeführt. Dadurch erhält man die quadratische Ungleichung

$$f(x) := x^2 - 2x \cos \omega T_d + 1 - \frac{4}{K} \sin^2 \omega T_d > 0. \quad (5.16)$$

Durch Lösen der quadratischen Gleichung $f(x) = 0$ kann obige Ungleichung umgeschrieben werden zu

$$(x - x_1)(x - x_2) > 0$$

mit

$$x_{1,2} = \cos \omega T_d \pm \sqrt{\frac{4}{K} - 1} \sin \omega T_d \quad (5.17)$$

Die Funktion $f(x)$ beschreibt eine nach oben offene Parabel. Besitzt sie konjugiert komplexe Nullstellen $x_{1,2}$, so liegt der Scheitelpunkt über der x -Achse. Damit ist die Ungleichung für alle $x \in \mathbb{R}$ erfüllt. Die Funktion $f(x)$ besitzt genau dann konjugiert komplexe Nullstellen, wenn gilt

$$K > 4. \quad (5.18)$$

D.h., für $K > 4$ bzw. $|b|^2 > 4|\lambda|^2$ ist μ_2 keine gültige Lösung, weshalb $\mu_C = \mu_1$ gilt.

Berücksichtigt man nun, dass $x = e^{-\sigma T_d}$ ein reeller und positiver Wert sein muss, besitzt die quadratische Gleichung nur dann eine gültige Lösung, wenn mindestens eine Lösung $x_{1,2}$ positiv ist. Sind beide Lösungen $x_{1,2}$ negativ, dann ist die Ungleichung (5.16) nicht erfüllt und nur μ_1 kann eine gültige Lösung für das Distanzmaß μ_C sein. Durch die Substitution

$$\cot \psi := \sqrt{\frac{4}{K} - 1} \quad 0 < \psi < \frac{\pi}{2}$$

kann die Beziehung (5.17) umgeschrieben werden zu

$$x_{1,2} = \frac{1}{\sin \psi} \sin(\psi \pm \omega T_d).$$

Damit beide Lösungen negativ sind, muss

$$\sin(\psi \pm \omega T_d) < 0$$

gelten. Diese Ungleichung ist im offenen Intervall

$$\omega T_d \in (\pi - \psi \quad \pi + \psi) \quad \text{mit} \quad \psi = \arcsin \sqrt{\frac{K}{4}} \quad (5.19)$$

erfüllt. D.h., für dieses Intervall gilt $\mu_C = \mu_1$.

Um die erhaltenen Ergebnisse zu verdeutlichen, wurde Ungleichung (5.15) für die Abtastzeiten

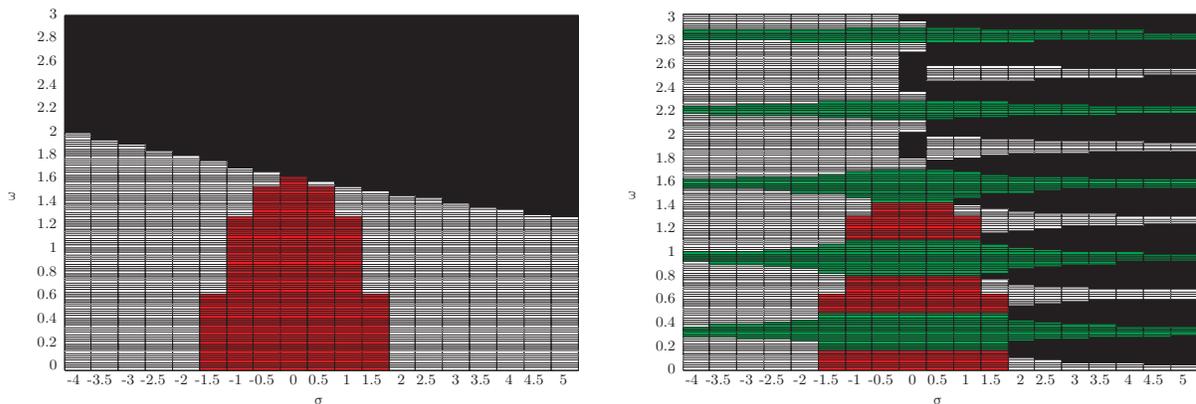


Abbildung 5.3: Graphische Darstellung der Ungleichung für $T_d = 0,1$ s (links) und $T_d = 10$ s (rechts)

$T_d = 0,1$ s und $T_d = 10$ s numerisch ausgewertet und über σ und ω in Abbildung 5.3 aufgetragen. Die Abbildung zeigt wie unterschiedlich sich die Ungleichung für kleine und große Abtastzeiten T_d verhält.

Die Wertepaare (σ, ω) , für welche die Ungleichung erfüllt ist, werden durch die schwarzen Flächen symbolisiert. Für diese Paare ist sowohl μ_1 als auch μ_2 eine mögliche Lösung für das Distanzmaß μ_C .

Die übrigen Flächen zeigen an, dass für diese Wertepaare (σ, ω) die Ungleichung nicht gilt, und somit nur μ_1 eine gültige Lösung sein kann. Die roten Flächen entsprechen den durch Gleichung (5.18) und die grünen Flächen den durch das Intervall (5.19) ermittelten Gültigkeitsbereich von μ_1 . Da das exakte Lösen der Ungleichung sehr aufwendig ist, können die weißen Flächen nicht durch eine analytische Lösung beschrieben werden.

Um weitere Aussagen über die Lösung der Ungleichung machen zu können, werden im Folgenden die Grenzwerte betrachtet. Untersucht man Ungleichung (5.15) für hinreichend große Abtastzeiten T_d , so kommt man zu dem Ergebnis, dass diese Ungleichung für $\sigma < 0$ nicht erfüllt ist. In diesem Fall gilt die Lösung μ_1 , und der Grenzwert von μ_1 bzw. μ geht für $T_d \rightarrow \infty$ gegen Null. Für hinreichend große T_d und $\sigma > 0$ gilt, dass im Intervall

$$\omega T_d \in (0 - \psi) \wedge (\pi - \psi - \pi + \psi) \wedge (2\pi - \psi - 2\pi) \quad \text{mit} \quad \psi = \arcsin \sqrt{\frac{K}{4}}$$

Lösung μ_1 gültig ist. Dieses Ergebnis stimmt mit der in Abbildung 5.3 dargestellten numerischen Auswertung überein.

Distanzmaß μ für die Abtastzeit $T_d = T_{d,K}$ Für Vielfache der Abtastzeit $T_{d,K}$ ist die Ungleichung (5.15) nicht erfüllbar. Somit ist für diese Abtastzeiten immer Lösung μ_1 gültig. Für sie gilt

$$\mu(kT_{d,K}) = \mu_1(kT_{d,K}) = 0 \quad \text{mit} \quad k = 0, 1, 2, \dots,$$

wodurch das Distanzmaß μ den Satz von KALMAN korrekt wiedergibt.

Ermittlung der optimalen Abtastzeit T_d^* Die Maxima der Lösung μ_1 bezüglich der Abtastzeit T_d können einfach ermittelt werden. Man erhält

$$T_{d,\max} = \frac{1}{\omega} \arctan\left(-\frac{\omega}{\sigma}\right) + \frac{k\pi}{\omega} \quad \text{mit } k = 0, 1, 2, \dots \quad (5.20)$$

Die Extrema der Lösung μ_2 konnten nicht analytisch ermittelt werden.

Spezialfall $\sigma = 0$ Für $\sigma = 0$ vereinfacht sich Ungleichung (5.15) zu

$$K - 2 \leq 2 \cos \omega T_d \quad \text{für } \omega T_d \neq k\pi.$$

Die Lösung μ_1 ergibt sich zu

$$\mu_1(\sigma = 0) = |\sin \omega T_d|$$

mit den Maxima

$$T_{d,\max} = (2k + 1) \frac{\pi}{2\omega} \quad \text{mit } k = 0, 1, 2, \dots$$

Für diesen Fall können auch die Extrema der Lösung (5.14)

$$\mu_2^2(\sigma = 0) = \frac{K(2 \sin^2 \omega T_d + K \cos \omega T_d - K)}{1 + \cos \omega T_d}$$

ermittelt werden. Da $T_{d,\max}$ nur positive, reelle Werte annehmen kann, erhält man für die möglichen Extrema

$$\cos \omega T_{d,\max} = -1 + \sqrt{K} \quad \text{mit } 0 < K < 1. \quad (5.21)$$

Es kann gezeigt werden, dass es sich bei diesem Extremum um ein Maximum der Lösung μ_2 handelt und dass es für das Intervall

$$K \in (0, 1)$$

dem Maximum des Distanzmaßes μ entspricht, da an der Stelle des Maximums gilt

$$\mu_1(T_{d,\max}) \geq \mu_2(T_{d,\max}).$$

Beispiel 5.1: Es wird nun das zeitkontinuierliche System

$$\mathbf{A} = \begin{bmatrix} \sigma & \omega \\ -\omega & \sigma \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

herangezogen. Für die Fälle $\sigma = 0$, $\sigma < 0$ und $\sigma > 0$ werden die Verläufe des Distanzmaßes μ untersucht und die optimale Abtastzeit T_d^* ermittelt. Die oben erarbeiteten analytischen Lösungen gelten nur für Systeme in Modalform. Das betrachtete System liegt nicht in Modalform vor. Die Ergebnisse können jedoch angewendet werden, da das System durch eine unitäre Transformation auf Modalform transformiert werden kann und das Distanzmaß invariant gegenüber dieser Transformation ist.

Die numerische Berechnung des Distanzmaßes erfolgt mittels Trisection-Algorithmus (siehe Abschnitt 3.2.1.4).

1. Fall $\sigma = 0$: Gegeben sei das zeitkontinuierliche System

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

mit den Eigenwerten $\lambda_{1,2} = \pm j$. Das Distanzmaß μ sowie die optimale Abtastzeit können *analytisch* ermittelt werden. Da gilt $K = \frac{|b|^2}{\sigma^2 + \omega^2} = 1$, ist das mit Hilfe der Lösung μ_1 ermittelte Maximum gültig und man erhält für die optimale Abtastzeit

$$T_d^* = (2k + 1) \frac{\pi}{2\omega} \quad \text{mit } k = 0, 1, 2, \dots$$

und den größtmöglichen Wert für das Distanzmaß $\mu(T_d^*) = 1$ (siehe Abbildung 5.4).

Wählt man hingegen als Eingangsvektor $\mathbf{b} = \left[\frac{1}{2} \quad \frac{1}{2}\right]^T$, dann gilt $K = \frac{1}{4}$. Um ein optimales T_d^* zu berechnen, muss das für die Lösung μ_2 geltende Maximum aus Gleichung (5.21) benutzt werden. Man erhält die Maxima

$$T_d^* = \arccos(-1 + \sqrt{K}) = 2,1 \text{ s} \quad \text{bzw.} \quad T_d^* = 2\pi - \arccos(-1 + \sqrt{K}) = 4,2 \text{ s}$$

mit dem Wert $\mu(T_d^*) = \frac{3}{4}$.

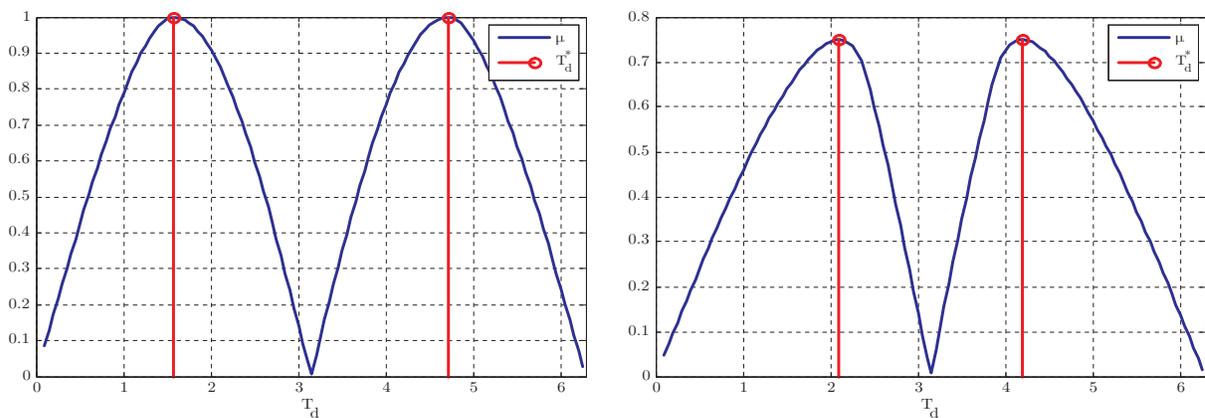


Abbildung 5.4: Distanzmaß $\mu(T_d)$ für $\sigma = 0$ und $\mathbf{b} = [1 \ 1]^T$ (links) und $\mathbf{b} = \left[\frac{1}{2} \quad \frac{1}{2}\right]^T$ (rechts)

2. Fall $\sigma < 0$: Für das zeitkontinuierliche System

$$\mathbf{A} = \begin{bmatrix} -1 & 1 \\ -1 & -1 \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

gilt $\lambda_{1,2} = -1 \pm j$ und $K = 1$. In obigem Abschnitt wurde gezeigt, dass für $K \leq 4$ der Gültigkeitsbereich der Lösungen μ_1 und μ_2 nicht für alle Abtastzeiten T_d bestimmt werden kann. Jedoch gilt in diesem Fall für hinreichend große Abtastzeiten Lösung μ_1 . Da für den Grenzwert von $\mu_1(T_d)$ gilt

$$\lim_{T_d \rightarrow \infty} \mu_1 = 0$$

und somit die Keulen zwischen den Nullstellen immer kleiner werden, wird angenommen, dass das Maximum T_d^* im Intervall $T_d^* \in [0 \ T_{d,K}]$ liegt, d.h. zwischen den ersten beiden Nullstellen von $\mu(T_d)$. Sucht man in diesem Intervall mit der BFGS-Methode nach einem Maximum, so erhält man die optimale Abtastzeit $T_d^* = 0,786 \text{ s}$ mit $\mu(T_d^*) = 0,322$. Im Intervall $[T_{d,K} \ 2T_{d,K}]$

befindet sich das Maximum an der Stelle $T_d = 3,927\text{ s}$ mit dem Wert $\mu = 0,014$ (vgl. Abbildung 5.5).

Für die analytisch ermittelbaren Maxima der Lösung μ_1 nach Gleichung (5.20) ergeben sich

$$T_{d,max} = 0,786\text{ s}, 3,927\text{ s}, \dots$$

In diesem Beispiel scheinen die Abtastzeiten $T_{d,max}$ hinreichend groß zu sein, so dass für diese Werte die Lösung μ_1 gilt und die maximale Abtastzeiten somit analytisch berechenbar wären. Jedoch kann vorab „hinreichend große Abtastzeit“ nicht näher spezifiziert werden.

3. Fall $\sigma > 0$: Für das kontinuierliche System

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

ist $\sigma = 0,5 > 0$ und $K = 1$. In diesem Fall kann auch für große T_d außerhalb des Intervalls (5.19) kein Gültigkeitsbereich angegeben werden, weshalb die Maxima bzw. die optimale Abtastzeit für $\sigma > 0$ und $K \leq 4$ *algorithmisch* bestimmt werden müssen. Die numerische Berechnung zeigt, dass die Keulen zwischen den Nullstellen von $\mu(T_d)$ mit größer werdenden Abtastzeiten höher werden. Man kann in einem Intervall zwischen zwei Nullstellen nach einer suboptimalen Abtastzeit suchen. Abbildung 5.5 zeigt die ersten beiden mit der BFGS-Methode gefundenen Maxima für die Abtastzeit $T_d = 2,17\text{ s}$ mit dem Wert $\mu = 2,43$ und $T_d = 5,46\text{ s}$ mit $\mu = 10,66$.

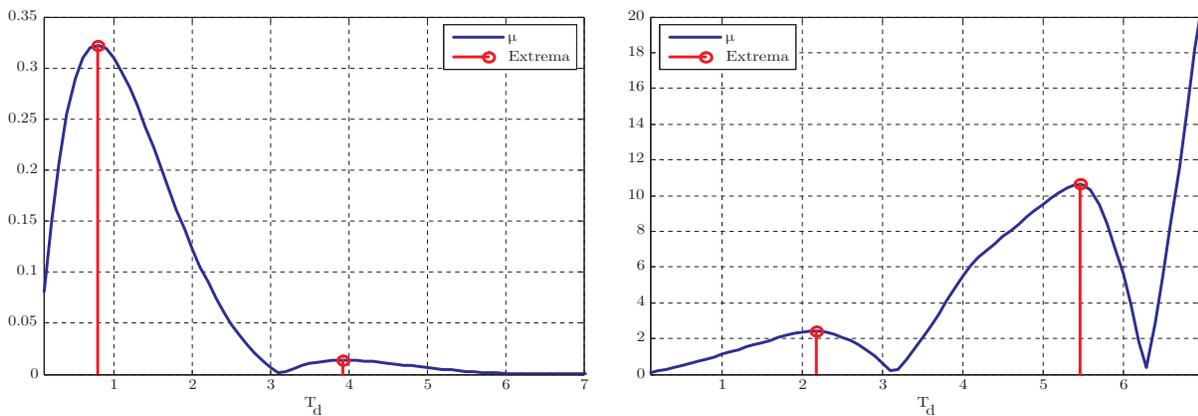


Abbildung 5.5: Distanzmaß μ für $\sigma < 0$ (links) und $\sigma > 0$ (rechts)

5.2.2.3 Vergleich der Maße

In dem nun folgenden Beispiel werden die hier untersuchten Maße κ nach LITZ sowie das Distanzmaß μ und die mit deren Hilfe ermittelten optimalen Abtastzeiten für reelle als auch komplexe Eigenwerte der Matrix \mathbf{A} verglichen.

Beispiel 5.2:

Reelle Eigenwerte: Gegeben sei das kontinuierliche System

$$\mathbf{A} = \begin{bmatrix} -2 & 0 \\ 0 & -1 \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

Abbildung 5.6 zeigt das Distanzmaß μ über der Abtastzeit T_d . Hier wurde die optimale Abtastzeit T_d^* mit Hilfe der BFGS-Methode ermittelt. Man erhält das Ergebnis $T_d^* = 0,65$ s und $\mu(T_d^*) = 0,12$. Die Maßfunktionen κ_i besitzen keine Maxima und streben für $T_d \rightarrow \infty$ einem endlichen Grenzwert zu.

Für eine gewünschte Mindest-Güte $\Gamma = 0,1$ errechnet sich nach Gleichung (5.12) aus dem Steuerbarkeitsmaßen κ_i , dass diese für $T_d > 0,5$ s gegeben ist. Mittels des Distanzmaßes μ kommt man zu dem Ergebnis, dass die Mindest-Güte im Intervall $T_d \in [0,33 \ 1,13]$ erreicht wird. Dieses Intervall kann nicht analytisch bestimmt werden und wurde aus Abbildung 5.6 abgelesen.

Man erhält zwei widersprüchliche Ergebnisse: Während κ anzeigt, dass mit steigender Abtastzeit die Steuerbarkeit zunimmt, kann durch Betrachten der Maßfunktion μ darauf geschlossen werden, dass die Steuerbarkeit mit steigender Abtastzeit wieder abnimmt.

Konjugiert komplexe Eigenwerte: Für das kontinuierliche System

$$\mathbf{A} = \begin{bmatrix} -1 & 1 \\ -1 & -1 \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

mit dem konjugiert komplexen Eigenwertpaar $\lambda_{1,2} = -1 \pm j$ wurde in Beispiel 5.1 das Distanzmaß μ berechnet. Mit Hilfe der BFGS-Methode wurde die optimale Abtastzeit $T_d^* = 0,786$ s bei $\mu(T_d^*) = 0,322$ ermittelt. In Abbildung 5.6 ist die Maßfunktion κ dargestellt. Für dieses Maß erhält man mit der BFGS-Methode die Werte $T_d^* = 2,29$ s und $\kappa(T_d^*) = 0,52$.

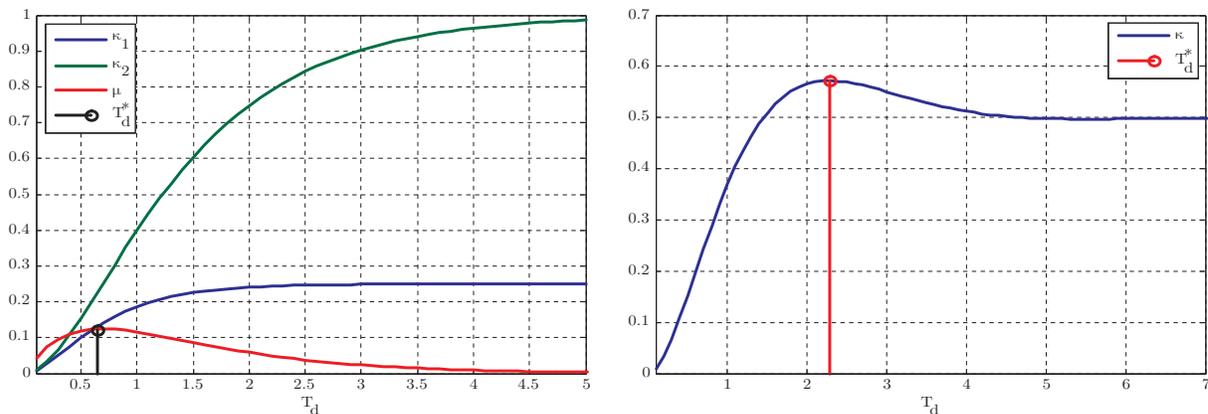


Abbildung 5.6: Maße für reelle Eigenwert (rechts) und Steuerbarkeitsmaß κ für konjugiert komplexe Eigenwerte (links); vgl. Distanzmaß in Abbildung 5.5 (rechts)

Auch dieses Beispiel zeigt, dass die unterschiedlichen Maßfunktionen zu sehr unterschiedlichen Ergebnissen führen, weshalb in Abschnitt 5.3 eine genauere Untersuchung anhand eines Zustandsreglerentwurfs durchgeführt wird.

Zusammenfassung (Diskreter Simulator für Systeme 2. Ordnung):

1. Die Maße streben für $T_d \rightarrow 0$ gegen Null.

2. Sind die Eigenwerte der Matrix \mathbf{A} reell ($\lambda_{1,2} \in \mathbb{R}$), können die ermittelten Ergebnisse wie folgt zusammengefasst werden:

		Grenzwerte	optimale Abtastzeit T_d^*
κ	$\lambda_{1,2} < 0$	$\lim_{T_d \rightarrow \infty} \kappa \leq c < \infty$	liegt im Unendlichen
	$\lambda_1 > 0 \vee \lambda_2 > 0$	$\lim_{T_d \rightarrow \infty} \kappa = \infty$	liegt im Unendlichen
μ	$\lambda_{1,2} < 0$	$\lim_{T_d \rightarrow \infty} \mu = 0$	algorithmisch ermittelbar
	$\lambda_1 > 0 \vee \lambda_2 > 0$	$\lim_{T_d \rightarrow \infty} \mu \leq c < \infty$	algorithmisch ermittelbar
	$\lambda_{1,2} > 0$	$\lim_{T_d \rightarrow \infty} \mu = \infty$	suboptimale T_d algorithmisch ermittelbar

Hierbei entspricht der Parameter c einer reellen und positiven Konstante.

3. Für konjugiert komplexe Eigenwerte $\lambda_{1,2} = \sigma \pm j\omega$ gilt:

		Grenzwerte	optimale Abtastzeit T_d^*	Abtastzeit $T_{d,K}$ (KALMAN)
κ	$\sigma < 0$	$\lim_{T_d \rightarrow \infty} \kappa < c < \infty$	algorithmisch ermittelbar	$\kappa(T_{d,K}) \neq 0$
	$\sigma = 0$	Funktion beschränkt	analytisch ermittelbar; System für T_d^* nicht steuerbar	$\kappa(T_{d,K}) = 0$ gilt nicht für alle Vielfache von $T_{d,K}$
	$\sigma > 0$	$\lim_{T_d \rightarrow \infty} \kappa = \infty$	liegt im Unendlichen	$\kappa(T_{d,K}) \neq 0$
μ	$\sigma < 0$	$\lim_{T_d \rightarrow \infty} \mu = 0$	für $K > 2$ analytisch, sonst algorithmisch ermittelbar	$\mu(T_{d,K}) = 0$
	$\sigma = 0$	Funktion beschränkt	analytisch berechenbar	$\mu(T_{d,K}) = 0$
	$\sigma > 0$	$\lim_{T_d \rightarrow \infty} \mu = \infty$	liegt im Unendlichen; suboptimale T_d für $K > 2$ analytisch, sonst algorithmisch ermittelbar	$\mu(T_{d,K}) = 0$

4. Die verschiedenen Maßzahlen liefern meist sehr unterschiedliche Ergebnisse. In Abschnitt 5.3 werden die beiden Maße anhand von Simulationen genauer untersucht, um festzustellen, welches der Maße ein gutes Ergebnis liefert.
5. Beim Steuerbarkeitsmaß κ ist aufgrund der Inkonsistenz mit dem Satz von KALMAN Vorsicht geboten.

5.2.3 Systeme höherer Ordnung

Steuerbarkeitsmaß κ Die reelle Systemmatrix \mathbf{A} besitzt entweder reelle Eigenwerte oder konjugiert komplexe Eigenwertpaare. Ausgehend von der Modalform werden die einzelnen Differentialgleichungen unabhängig voneinander diskretisiert, wodurch die Diagonalstruktur erhalten bleibt.

Für das Steuerbarkeitsmaß κ erhält man für reelle Eigenwerte die in Abschnitt 5.2.1.1 und für konjugiert komplexe Eigenwerte die in Abschnitt 5.2.2.1 untersuchten Maßfunktionen in Abhängigkeit der Abtastzeit T_d . Durch die Analyse des diskreten Simulators 1. und 2. Ordnung können folgende Aussagen über Systeme höherer Ordnung getroffen werden: Besitzt die Systemmatrix \mathbf{A} nur reelle Eigenwerte λ_i ($i = 1, \dots, n$), so erhält man n monoton steigende Maßfunktionen κ_i . Somit gilt, dass mit steigender Abtastzeit die Steuerbarkeit jedes Eigenwertes zunimmt. In diesem Fall kann die Frage nach der Mindest-Güte bei gegebener Abtastzeit bzw.

nach der minimalen Abtastzeit bei einer gewünschten Mindest-Güte einfach auf Systeme höherer Ordnung erweitert werden.

Für konjugiert komplexe Eigenwerte der Matrix \mathbf{A} ist das Steuerbarkeitsmaß κ für die Beurteilung der Steuerbarkeit nicht anwendbar, da die durch die Maßfunktion erhaltenen Ergebnisse nicht mit dem Satz von KALMAN (Theorem 12) konsistent sind. D.h. sie zeigen den Steuerbarkeitsverlust für Vielfache der Abtastzeit $T_d = kT_{d,K}$ ($k = 1, 2, \dots$) nicht korrekt an (siehe Abschnitt 5.2.2.1).

Distanzmaß μ Für die Berechnung des Distanzmaßes μ wird das Gesamtsystem betrachtet, wodurch eine Verallgemeinerung der bei Systemen 1. und 2. Ordnung erhaltenen Ergebnisse nicht möglich ist. Da schon bei Systemen 2. Ordnung für die meisten Fälle keine analytische Lösung gefunden werden kann, ist zu erwarten, dass dies auch bei Systemen höherer Ordnung nicht möglich ist.

Es hat sich jedoch gezeigt, dass die numerische Berechnung mittels Trisection-Algorithmus in Verbindung mit GUs Test und die Ermittlung der optimalen Abtastzeit T_d^* mit Hilfe der BFGS-Methode i.A. zufrieden stellende Ergebnisse liefern. Schwierigkeiten bei der Bestimmung von T_d^* können vor allem dann auftreten, wenn die Maßfunktion mehrere lokale Minima aufweist. Da die BFGS-Methode ein lokales Minimum liefert, ist das Ergebnis stark von dem vorgegebenen Startwert abhängig.

5.3 Anwendung und Deutung der Maße

Es hat sich gezeigt, dass die hier untersuchten Maße unterschiedliche Verläufe und auch unterschiedliche Ergebnisse für die optimale Abtastzeit T_d^* liefern. In diesem Abschnitt wird versucht, die erhaltenen Ergebnisse durch Einsatz eines diskreten Zustandsreglers zu überprüfen bzw. zu deuten. Es wird untersucht, welchen Einfluss der durch die Abtastzeit festgelegte Wert des Steuerbarkeitsmaßes der Strecke auf den Regelkreis hat. Hierfür werden der Verlauf der Maße in Abhängigkeit der Abtastzeit T_d sowie die optimale Abtastzeit T_d^* mit einfach berechenbaren bzw. messbaren und physikalisch aussagekräftigen Größen verglichen. Für die praktische Realisierung sind bei der Wahl der Abtastzeit weitere physikalische Einschränkungen zu berücksichtigen, welche hier außer Acht gelassen werden. Jedoch liefert die durchgeführte theoretische Untersuchung eine Hilfestellung bei der Wahl einer „guten“ Abtastzeit bezüglich eines Steuerbarkeitsmaßes innerhalb dieser Beschränkungen. Für die Untersuchung werden folgende Größen herangezogen:

Verstärkung: Die „Verstärkung“ des Zustandsreglers (siehe Abschnitt 5.3.1) wird als Anzeichen für „gute“ bzw. „schlechte“ Steuerbarkeit herangezogen. Intuitiv kann man sich leicht vorstellen, dass für eine schlecht steuerbare Strecke die Zustandsvariablen x stärker gewichtet („verstärkt“) werden müssen, um eine dementsprechend hohe Stellgröße u zu erzeugen. D.h., es ist anzunehmen, dass in diesem Fall die Verstärkung betragsmäßig hohe Werte annimmt.

Stelleistung: Ein weiteres Anzeichen für die „Güte“ der Steuerbarkeit ist die Leistung der Stellgröße u , da je schlechter ein System steuerbar ist, umso mehr Leistung wird benötigt, um das System zu „steuern“.

5.3.1 Zustandsregler

Für das zeitdiskrete, steuerbare System der Form

$$\boldsymbol{\xi}_{i+1} = \mathbf{A}_d \boldsymbol{\xi}_i + \mathbf{b}_d u_i \quad \eta_i = \mathbf{c}_d^T \boldsymbol{\xi}_i.$$

wird ein Zustandsregler mit dem Regelgesetz

$$u_i = -\mathbf{h}^T \boldsymbol{\xi}_i$$

entworfen. Für den geschlossenen Regelkreis ergibt sich die Beschreibung

$$\boldsymbol{\xi}_{i+1} = (\mathbf{A}_d - \mathbf{b}_d \mathbf{h}^T) \boldsymbol{\xi}_i = \tilde{\mathbf{A}}_d \boldsymbol{\xi}_i \quad \eta_i = \mathbf{c}_d^T \boldsymbol{\xi}_i.$$

Durch Vorgabe der gewünschten Eigenwerte $\tilde{\lambda}_i$ ($i = 1, \dots, n$) der Matrix $\tilde{\mathbf{A}}_d$ des geschlossenen Regelkreises kann der Vektor \mathbf{h}^T bestimmt werden.

Für ein in Modalform vorliegendes System

$$\hat{\mathbf{A}}_d = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{bmatrix} \quad \hat{\mathbf{b}}_d = \begin{bmatrix} \hat{b}_{d,1} \\ \hat{b}_{d,2} \\ \vdots \\ \hat{b}_{d,n} \end{bmatrix}$$

kann folgende explizite Relation für die gesuchten Werte h_i in Abhängigkeit der Eigenwerte λ_i und der Wunscheigenwerte $\tilde{\lambda}_i$ angegeben werden:

$$h_i = \frac{1}{\hat{b}_{d,i}} \frac{\prod_{j=1}^n (\lambda_i - \tilde{\lambda}_j)}{\prod_{j=1, j \neq i}^n (\lambda_i - \lambda_j)} \quad i = 1, 2, \dots, n.$$

Hieraus lässt sich der Zusammenhang zwischen den Modalmaßen und den Werten h_i erkennen. Wenn der Wert $\hat{b}_{d,i}$ klein ist, erhält man kleine Werte für die Modalmaße und der zugehörige Wert h_i ist groß.

Die für die Untersuchung benötigte Verstärkung ist durch Bilden der Quadratsumme der einzelnen Elemente des n dimensionalen Vektors \mathbf{h}

$$H := \mathbf{h}^T \mathbf{h}$$

definiert. Für die Bestimmung der Leistung wird der Regelkreis simuliert. Der aus der Simulation erhaltenen Werte u_i werden durch der Anzahl der Stützstellen N normiert, wodurch man die Größe

$$U := \frac{1}{N} \sum_{i=0}^N u_i^2$$

erhält.

Die beiden Größen H und U werden für verschiedene Abtastzeiten T_d ermittelt. Da sowohl die Funktion $H(T_d)$ als auch $U(T_d)$ von den vorgegebenen Eigenwerten $\tilde{\lambda}_i$ der Matrix $\mathbf{A}_d - \mathbf{b}_d \mathbf{h}^T$ abhängig ist, werden diese Funktionen für unterschiedliche Werte von $\tilde{\lambda}_i$ ermittelt. Die vorgegebenen Wunscheigenwerte $\tilde{\lambda}$ sind gleichmäßig im Einheitskreis verteilt (vgl. Abbildung 5.7).

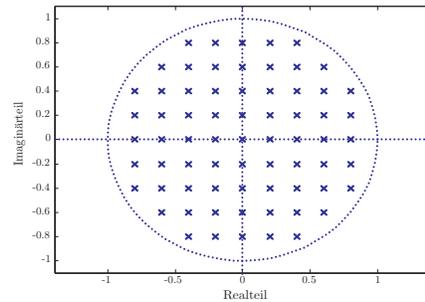


Abbildung 5.7: Wunscheigenwerte $\tilde{\lambda}_i$

5.3.1.1 Systeme 1. Ordnung

Für einen aus dem zeitkontinuierlichen System

$$\frac{dx}{dt} = ax + du$$

entstandenen zeitdiskreten Simulator 1. Ordnung

$$\zeta_{i+1} = a_d \zeta_i + b_d u_i$$

können die Verstärkung h und die Stellgröße u einfach berechnet werden. Man erhält

$$h = \frac{a_d - \tilde{\lambda}}{b_d} \quad \text{und} \quad u_i = \frac{(\tilde{\lambda} - a_d) \tilde{\lambda}^i x_0}{b_d}$$

bzw.

$$h = \frac{a (e^{aT_d} - \tilde{\lambda})}{b (e^{aT_d} - 1)} \quad \text{und} \quad u_i = \frac{a (\tilde{\lambda} - e^{aT_d})}{b (e^{aT_d} - 1)} \tilde{\lambda}^i x_0.$$

Die Größe $H = h^2$ strebt für $T_d \rightarrow 0$ gegen Unendlich. Für $T_d \rightarrow \infty$ strebt sie dem Grenzwert

$$\lim_{T_d \rightarrow \infty} H(T_d) = \begin{cases} \left(\frac{\tilde{\lambda}a}{b}\right)^2 & \text{für } a < 0 \\ \left(\frac{a}{b}\right)^2 & \text{für } a > 0 \end{cases}$$

zu. Die Funktion $H(T_d)$ hat für die Abtastzeit

$$T_{d,\min} = \frac{\ln \tilde{\lambda}}{a}$$

ein Minimum, für das gilt $H(T_{d,\min}) = 0$. Ist $T_{d,\min}$ nicht reell und positiv, dann nimmt die Funktion $H(T_d)$ mit zunehmender Abtastzeit ab und zeigt somit an, dass die Steuerbarkeit mit steigendem T_d zunimmt. Das stimmt mit den aus den monoton steigenden Maßfunktionen κ und μ erhaltenen Ergebnissen überein.

Die oben erwähnte Nullstelle könnte als die Abtastzeit mit bestmöglicher Steuerbarkeit gedeutet werden, da keine Stellgröße benötigt wird, um den Systemzustand in den Ursprung zu überführen. Dies wird von den Maßen nicht wiedergegeben, was sich wie folgt erklären lässt: Die Abtastzeit $T_{d,\min}$ entspricht genau der Abtastzeit, für die der gewünschte Eigenwert $\tilde{\lambda}$ dem

Eigenwert des unregulierten Systems entspricht ($\tilde{\lambda} = e^{aT_d, \min}$). Dadurch ist kein Eingriff des Zustandsreglers notwendig, und damit ist die Verstärkung gleich Null. Die Maße zeigen dies nicht an, da sie unabhängig von der Wahl der Wunscheigenwerte sind und dieser Effekt unabhängig von der Güte der Steuerbarkeit des Systems ist.

Für die Funktion $U(T_d)$ erhält man

$$U = \frac{1}{N} \sum_{i=0}^N u_i^2 = \frac{x_0^2 a^2 (\tilde{\lambda} - e^{aT_d})^2 ((\tilde{\lambda}^2)^N \tilde{\lambda}^2 - 1)}{Nb^2 (e^{aT_d} - 1)^2 (\tilde{\lambda}^2 - 1)}.$$

Die Funktion strebt für $T_d \rightarrow 0$ gegen Unendlich und für $T_d \rightarrow \infty$ gegen die Grenzwerte:

$$a < 0 : \quad \frac{\tilde{\lambda}^2 x_0^2 a^2 ((\tilde{\lambda}^2)^N \tilde{\lambda}^2 - 1)}{Nb^2 (\tilde{\lambda}^2 - 1)}$$

$$a > 0 : \quad \frac{x_0^2 a^2 ((\tilde{\lambda}^2)^N \tilde{\lambda}^2 - 1)}{Nb^2 (\tilde{\lambda}^2 - 1)}.$$

Aus der Funktion $U(T_d)$ können die gleichen Schlussfolgerungen bezüglich der Monotonieeigenschaften und der Nullstelle wie bei der Funktion $H(T_d)$ gezogen werden.

5.3.1.2 Systeme 2. Ordnung

Für Systeme 2. Ordnung wurden die Größen $H(T_d)$ und $U(T_d)$ nicht analytisch ermittelt, sondern für die verschiedenen Wunscheigenwerte mit Hilfe von Matlab/Simulink numerisch berechnet. Im Folgenden werden die Ergebnisse anhand zweier Beispiele erläutert.

Reelle Eigenwerte: Gegeben sei das schon in Beispiel 5.2 untersuchte System

$$\mathbf{A} = \begin{bmatrix} -2 & 0 \\ 0 & -1 \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

mit den reellen Eigenwerten $\lambda_1 = -2$ und $\lambda_2 = -1$. Bei der in Abschnitt 5.2.2 durchgeführten Untersuchung des diskreten Simulators 2. Ordnung wurde festgestellt, dass beide Maße für $T_d \rightarrow 0$ gegen Null streben. Dies lässt auch Abbildung 5.8 erkennen, da für $T_d \rightarrow 0$ die Werte von $H(T_d)$ und $U(T_d)$ ansteigen. Aus dieser Abbildung ist auch ersichtlich, dass beide Größen mit steigender Abtastzeit T_d zunehmen. D.h., die Steuerbarkeit nimmt mit steigender Abtastzeit ab. Diesen Sachverhalt gibt nur das Distanzmaß μ korrekt wieder.

Aus Abbildung 5.8 sieht man, dass die optimale Abtastzeit T_d^* abhängig von der Wahl der Eigenwerte $\tilde{\lambda}$ ist. Für den Großteil der vorgegebenen Eigenwerte liegt die optimale Abtastzeit im Intervall $T_d^* \in [0, 2]$. In Tabelle 5.1 sind die durch die Maße erhaltenen optimalen Abtastzeiten aufgelistet.

Wie aus der Tabelle ersichtlich ist, liefert nur das Distanzmaß μ eine optimale Abtastzeit, für die $T_d^* \in [0, 2]$ gilt.

Konjugiert komplexe Eigenwerte: Gegeben sei das bereits in den Beispielen 5.1 und 5.2 untersuchte System

$$\mathbf{A} = \begin{bmatrix} -1 & 1 \\ -1 & -1 \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

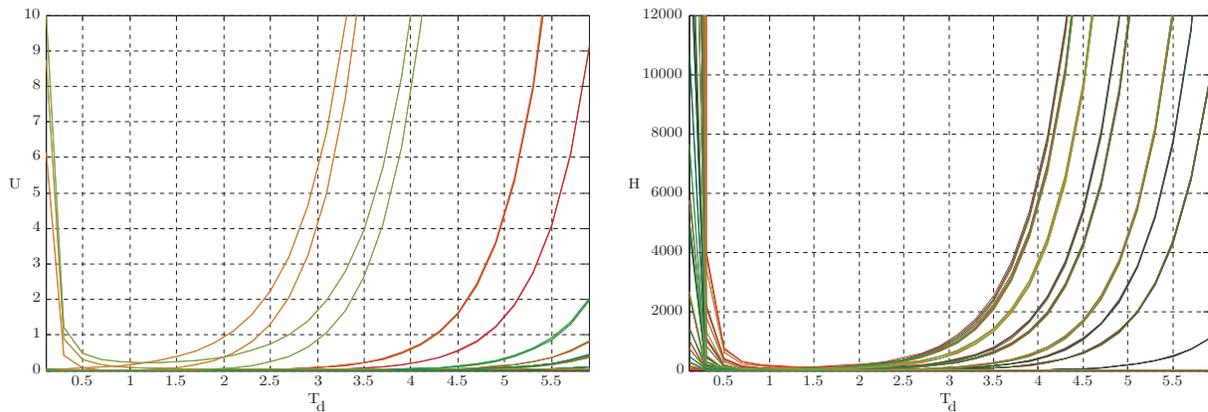


Abbildung 5.8: Diskreter Simulator: Leistung der Stellgröße U (links) und Verstärkung H (rechts)

Tabelle 5.1: optimale Abtastzeiten (diskreter Simulator)

Maß	T_d^* [s]	$T_d^* \in [0,2 \ 2]$
μ	0,65	✓
κ	∞	✗

mit den konjugiert komplexen Eigenwerten $\lambda_{1,2} = 1 \pm j$. In Abbildung 5.9 sind für den diskreten Simulator die Größen U und H in Abhängigkeit der Abtastzeit T_d für verschiedene Werte von $\tilde{\lambda}_i$ dargestellt.

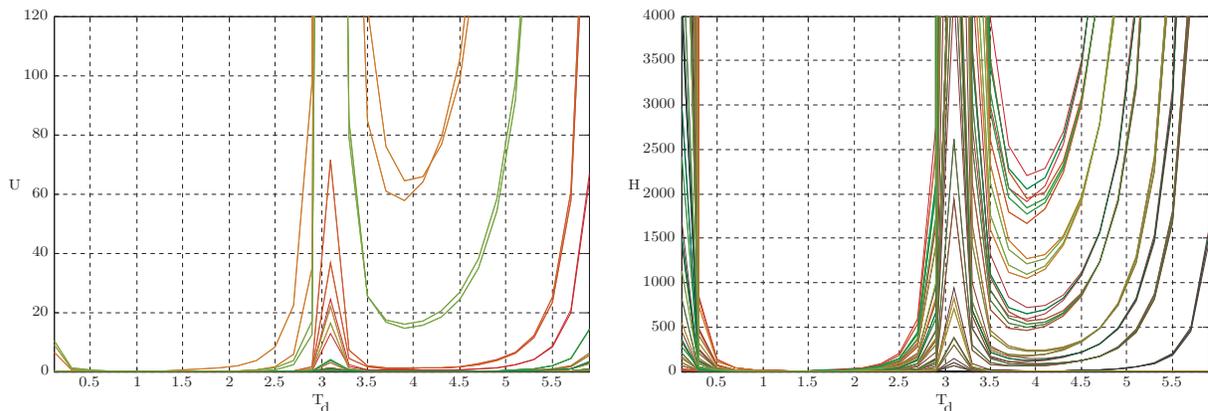


Abbildung 5.9: Diskreter Simulator: Leistung der Stellgröße U (links) und Verstärkung H (rechts)

Laut Theorem 12 nach KALMAN ist das System für Vielfache der Abtastzeit

$$T_{d,K} = \frac{\pi}{\omega}$$

nicht steuerbar. Die Größe H nimmt in der Umgebung dieser Abtastzeiten ($T_{d,K} = \frac{\pi}{\omega} = \pi$) sehr große Werte an, was auf schlechte Steuerbarkeit hinweist, während in den Intervallen zwischen diesen Abtastzeiten gute Steuerbarkeit angezeigt wird. Weiters lässt die Funktion $H(T_d)$ erkennen, dass für das hier betrachtete System mit steigender Abtastzeit die Steuerbarkeit in diesen

Intervallen abnimmt und die optimale Abtastzeit im Intervall $T_d^* \in [0 \pi]$ liegt. Hierbei ist die genaue Position des Minimums abhängig von den gewünschten Eigenwerten $\tilde{\lambda}_i$ des geschlossenen Kreises. Die Betrachtung der Leistung der Stellgröße U lässt die gleichen Schlussfolgerungen zu. Genau dieser Sachverhalt wird von der Funktion $\mu(T_d)$ in Abbildung 5.5 wiedergegeben. Für dieses Beispiel liefert das Distanzmaß μ die optimale Abtastzeit $T_d^* = 0,786$ im Intervall $[0 \pi]$.

Auf die Untersuchung des Steuerbarkeitsmaßes κ wird hier verzichtet. Bereits die vorangegangenen Untersuchungen haben gezeigt, dass es für zeitdiskrete Systeme mit konjugiert komplexen Eigenwerten nicht geeignet erscheint, da das Maß den Steuerbarkeitsverlust für Vielfache der Abtastzeit $T_{d,K}$ nicht anzeigt und somit das Theorem nach KALMAN nicht wiedergibt. Dies liegt daran, dass der Steuerbarkeitsverlust für Vielfache dieser Abtastzeit durch einen doppelten Eigenwert hervorgerufen wird und die Modalmaße für diesen Fall genau genommen nicht definiert sind.

5.3.2 Anwendbarkeit

In diesem Abschnitt werden analoge Untersuchungen mittels Zustandsregler für die schon in Kapitel 4.2 benutzten Beispiele „Wagen mit Stab“, „Hubmagnet“ und „Doppelpendel“ durchgeführt. Da auf die Modellbildung und die Linearisierung um einen Arbeitspunkt bereits ausführlich eingegangen wurde, wird hier auf genaue Ausführungen verzichtet.

Wagen mit Stab:

Für den in Abbildung 5.10 dargestellten Wagen mit Stab wurde in Abschnitt 4.2.1 ein nichtlineares Zustandsraummodell ermittelt. Mit den Zustandsvariablen

$$x_1 := \varphi, \quad x_2 := \frac{d\varphi}{dt}$$

erhält man das um die Ruhelage $\mathbf{x}_R = \mathbf{0}$ und $u_R = 0$ linearisierte System

$$\mathbf{A} := \left. \frac{\partial \mathbf{f}}{\partial \mathbf{x}} \right|_{\mathbf{x}=\mathbf{x}_R, u=u_R} = \begin{bmatrix} 0 & 1 \\ g & 0 \end{bmatrix} \quad \text{und}$$

$$\mathbf{b} := \left. \frac{\partial \mathbf{f}}{\partial u} \right|_{\mathbf{x}=\mathbf{x}_R, u=u_R} = \begin{bmatrix} 0 \\ -1 \end{bmatrix}.$$

Die Matrix \mathbf{A} besitzt die reellen Eigenwerte

$$\lambda_{1,2} = \pm \sqrt{g}.$$

Da einer der beiden Eigenwerte positiv ist, kann man auf Grund der Ergebnisse der untersuchten Maßzahlen schließen, dass mit steigender Abtastzeit T_d die Steuerbarkeit des diskreten Simulators zunimmt. Dieses Resultat kann durch die aus dem Zustandsreglerentwurf ermittelte Größe H und der Stelleistung U bestätigt werden.

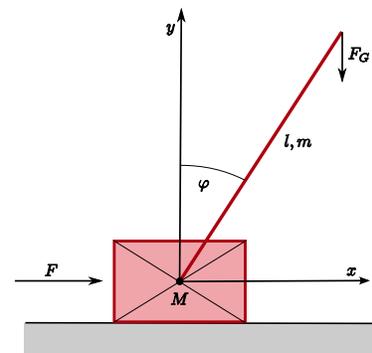


Abbildung 5.10: Wagen mit Stab

Hubmagnet:

Als zweites Beispiel werden die Steuerbarkeitsmaße des in Abbildung 5.11 dargestellten Hubmagneten für verschiedene Kugelmassen untersucht. Für dieses Labormodell wurde in Abschnitt 4.2.2 das um die Ruhelage

$$\mathbf{x}_R = \left[y_R, 0, \sqrt{\frac{Mg}{c}} y_R \right]^T, \quad u_R = R y_R \sqrt{\frac{Mg}{c}}$$

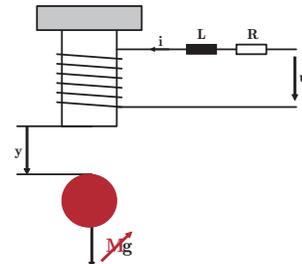


Abbildung 5.11: Hubmagnet

linearisierte Modell

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 \\ \frac{2g}{y_R} & 0 & -\frac{2}{y_R} \sqrt{\frac{cg}{M}} \\ 0 & \frac{2}{Ly_R} \sqrt{cMg} & -\frac{1}{L} \end{bmatrix} \quad \text{und} \quad \mathbf{b} = \begin{bmatrix} 0 \\ 0 \\ \frac{1}{L} \end{bmatrix}$$

ermittelt. Das System besitzt drei reelle Eigenwerte, von denen einer positiv ist. Wir wissen, dass das zu diesem positiven Eigenwert gehörende Steuerbarkeitsmaß κ für $T_d \rightarrow \infty$ gegen Unendlich strebt. Für das Distanzmaß μ ist für dieses System keine analytische Lösung vorhanden, jedoch zeigt die numerische Lösung im Gegensatz zum Steuerbarkeitsmaß κ , dass die Steuerbarkeit mit steigender Abtastzeit abnimmt.

Um diese Ergebnisse genauer zu untersuchen, wurden folgende Experimente durchgeführt (siehe [8]): Für das linearisierte System wurden für unterschiedliche Abtastzeiten und Wunscheigenwerte Zustandsregler entworfen und mit dem linearisierten Modell simuliert. Durch die aus dieser Simulation ermittelten Funktionen $H(T_d)$ und $U(T_d)$ wurde der durch das Distanzmaß berechnete Steuerbarkeitsverlauf bestätigt. Weiters wurde die Funktion $U(T_d)$ an der realen Strecke ermittelt. Dadurch konnte folgendes festgestellt werden: Die Leistung U wird durch mehrere Faktoren beeinflusst, die durch das Distanzmaß nicht wiedergegeben werden. Diese sind unter anderem die Nichtlinearität des Systems, die starke zeitliche Veränderung des Spulenwiderstands durch Erwärmung sowie Windup Effekte. Durch die Regelung mit einem integrierenden Zustandsregler mit Anti-Windup nach HIPPE wurden die Auswirkungen einiger dieser Faktoren reduziert und ein der Simulation ähnliches Verhalten des Verlaufs der Leistung erzielt.

Doppelpendel:

Für das Doppelpendel gemäß Abbildung 5.12 wurde in Abschnitt 4.2.3 durch Linearisierung um die instabile Ruhelage $\mathbf{x}_R = \mathbf{0}$ und durch Einführen der Konstanten

$$M_1 = 3m_2m_3 + m_1m_2 + 4m_1^2 + 16m_1m_3,$$

$$M_2 = 5m_1m_2 + 4m_1^2 + m_2^2,$$

$$M_3 = 3m_2m_3 + 4m_1m_3 + m_1m_2 + m_1^2 \quad \text{und}$$

$$M_4 = 3m_2m_3 + m_1m_2 + 4m_1^2 + 16m_1m_3$$

das System

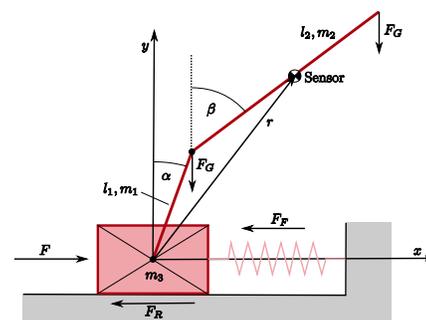


Abbildung 5.12: Doppelpendel

$$\mathbf{A} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ \frac{3g(M_2+4m_2m_3+4m_1m_3)}{l_1M_1} & -\frac{9m_3m_2g}{l_1M_1} & \frac{3c(4m_1+m_2)}{l_1M_1} & 0 & 0 & \frac{3k(4m_1+m_2)}{l_1M_1} \\ -\frac{9m_3g(m_2+m_1)}{l_2M_1} & \frac{3gM_3}{l_2M_1} & \frac{3m_1c}{l_2M_1} & 0 & 0 & \frac{3m_1k}{l_2M_1} \\ -\frac{3gM_2}{M_1} & -\frac{3m_1m_2g}{M_1} & -\frac{c(16m_1+3m_2)}{M_1} & 0 & 0 & -\frac{k(16m_1+3m_2)}{M_1} \end{bmatrix},$$

$$\mathbf{b} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ -\frac{3(m_2+4m_1)}{l_1M_4} \\ -\frac{3m_1}{l_2M_4} \\ \frac{16m_1+3m_2}{M_4} \end{bmatrix} \quad \text{und} \quad \mathbf{C} = \mathbf{E}.$$

ermittelt. Für die in Anhang B angeführten Parameterwerte besitzt das System ein konjugiert komplexes Eigenwertpaar

$$\lambda_{5,6} = -0,42 \pm j0,79.$$

Laut dem Satz von KALMAN ist der diskrete Simulator für die Abtastzeiten

$$\begin{aligned} T_{d,K} &= k \frac{\pi}{0,79} \quad \text{mit } k = 1, 2, 3, \dots \\ &= \{3,99; 7,98; \dots\} \end{aligned}$$

nicht steuerbar. Keines der Steuerbarkeitsmaße κ_i wird für die oben angeführten Werte von $T_{d,K}$ zu Null, weshalb der Steuerbarkeitsverlust nicht korrekt wiedergegeben wird. Für dieses Beispiel ist auch das numerisch ermittelte Distanzmaß nicht konsistent zu den KALMANschen Begriffen. Dies liegt daran, dass die Systemmatrix \mathbf{A} positive Eigenwerte besitzt, wodurch die Elemente der Matrix $\mathbf{A}_d = e^{\mathbf{A}T_d}$ mit größer werdender Abtastzeit sehr große Werte annehmen. Dadurch sind die numerischen Fehler sehr groß.

Für Abtastzeiten $T_d > 1,6$ liefert auch die numerische Auswertung des KALMAN-Kriteriums das falsche Ergebnis, dass das diskretisierte System nicht steuerbar sei.

5.4 Zusammenfassung

In diesem Kapitel wurde für den diskreten Simulator der Einfluss der Abtastzeit T_d auf die Steuerbarkeit untersucht und sofern möglich die optimale Abtastzeit T_d^* analytisch ermittelt.

Für den diskreten Simulator können die optimalen Abtastzeiten für Systeme 1. und 2. Ordnung in den meisten Fällen analytisch berechnet werden. Für ein konjugiert komplexes Eigenwertpaar der Matrix \mathbf{A} gibt nur das Distanzmaß μ den durch den Satz von KALMAN beschriebenen Steuerbarkeitsverlust für Vielfache der Abtastzeit $T_{d,K}$ korrekt wieder. Da diese Diskretisierungsmethode das System „strukturertretend“ diskretisiert, können auch für Systeme höherer

Ordnung die Maßfunktionen bzw. die optimalen Abtastzeiten für das Modalmaß κ für die einzelnen Eigenwerte direkt angegeben werden. Um die optimale Abtastzeit für das Gesamtsystem zu ermitteln, müssen entweder die einzelnen Funktionen geeignet zusammengefasst oder das Distanzmaß μ verwendet werden. In beiden Fällen kann eine analytische Lösung nicht angegeben werden.

Die Maße liefern meist sehr unterschiedliche Ergebnisse, weshalb die ermittelten optimalen Abtastzeiten mit den durch Simulation erhaltenen Größen H und U verglichen wurden. Betrachtet man die Beispiele aus Abschnitt 5.3.1.2 erkennt man, dass nur das Distanzmaß mit der Simulation übereinstimmende Ergebnisse liefert. Nur das Distanzmaß zeigt für den diskreten Simulator im Falle konjugiert komplexer Eigenwerte den Steuerbarkeitsverlust laut Theorem nach KALMAN korrekt an.

Mit Hilfe der hier erhaltenen Ergebnisse können meist nur für Systeme 1. und 2. Ordnung oder für diskretisierte Systeme, dessen zugehöriges kontinuierliches System Eigenwerte mit positivem Realteil besitzt, Aussagen über die optimale Abtastzeit T_d^* getroffen werden, ohne die Maßfunktionen berechnen bzw. die optimale Abtastzeit algorithmisch ermitteln zu müssen. Bei der algorithmischen Bestimmung der optimalen Abtastzeit T_d^* hat sich die BFGS-Methode als ein gutes Werkzeug erwiesen. Die Simulationen zeigen, dass für einen Zustandsregler die optimale Abtastzeit auch von den Eigenwerten des geschlossenen Kreises abhängig ist, die von den Maßen jedoch nicht berücksichtigt werden. Aus diesem Grund kann die durch die Maße ermittelte optimale Abtastzeit nicht unmittelbar mittels den Funktionen $U(T_d)$ und $H(T_d)$ überprüft werden.

Die Anwendungsbeispiele haben gezeigt, dass das Distanzmaß μ in den meisten Fällen Ergebnisse liefert, die mit den aus der Simulation erhaltenen Werten H und U besser übereinstimmen als die der Steuerbarkeitsmaße κ_i nach LITZ.

Zusammenfassend kann gesagt werden, dass bei der Wahl der Abtastzeit einige Aspekte zu berücksichtigen sind. Diverse physikalische und technische Einschränkungen (Hardware, Aliasing Effekt, Berechnungsdauer des Regelgesetzes, ...) müssen beachtet werden. In den hier durchgeführten Experimenten wurde die Abtastzeit bezüglich dem maximalen Wert von Steuerbarkeitsmaßen optimiert. Analog ist natürlich auch eine Optimierung bezüglich des Beobachtbarkeitsmaßes denkbar. Damit stellt sich die Frage, ob bezüglich der Steuerbarkeit, der Beobachtbarkeit oder einer Kombination der beiden Eigenschaften optimiert werden soll.

Für die Ermittlung der optimalen Abtastzeit des Gesamtsystems erweist sich das Distanzmaß μ als das Effizienteste der hier untersuchten Maßzahlen. Das Maß μ ist konsistent mit dem Steuerbarkeitsverlust nach KALMAN und kann z.B. mit dem Trisection-Algorithmus in Verbindung mit GUs Test zuverlässig ermittelt werden. Das Distanzmaß hat den Nachteil, dass in den meisten Fällen keine analytische Lösung für das Maß und somit auch für die optimale Abtastzeit vorhanden ist. Jedoch liefert die Maximumsuche der Funktion $\mu(T_d)$ mittels der BFGS meist gute Ergebnisse.

Auf Grund der strukturhaltenden Transformation können die Steuerbarkeitsmaße κ_i nach LITZ für die einzelnen Systemteile auch für Systeme höherer Ordnung analytisch ermittelt werden. Dadurch kann das Steuerbarkeitsverhalten bzw. die optimale Abtastzeit für einen bestimmten Eigenwert sehr einfach ermittelt werden. Bei diesem Maß muss jedoch beachtet wer-

den, dass nicht für alle Vielfache der Abtastzeit $T_{d,K}$ der Steuerbarkeitsverlust korrekt angezeigt wird.

Kapitel 6

Zusammenfassung und Ausblick

Im Allgemeinen ist festzuhalten, dass alle Maße Vor- und Nachteile besitzen und dass bei bestimmten Systemen Schwierigkeiten auftreten oder falsche Ergebnisse erhalten werden können. Aus diesem Grund ist es ratsam, sich die Struktur des Systems genau anzusehen, bevor man sich für ein bestimmtes Steuerbarkeitsmaß entscheidet. Meist gibt eine konkrete Anwendung die am ehesten zu benutzende Maßklasse bereits vor (siehe Abschnitt 4.3).

Die Modalmaße zeichnen sich durch ihre einfache Berechenbarkeit aus. Sie sollten jedoch nur im Fall von einfachen Eigenwerten angewendet werden, da sie bei einem Steuerbarkeitsverlust auf Grund von inneren Verkopplungen hervorgerufen durch mehrfache Eigenwerte keine konsistenten Ergebnisse liefern. Die fehlende Konsistenz macht sich jedoch auch bei nahe aneinander liegenden Eigenwerten bemerkbar, da in diesem Fall hohe Maßzahlen eine gute Steuerbarkeit vortäuschen. Es lässt sich leider keine generelle Aussage über den notwendigen Abstand der Eigenwerte treffen, ab dem die Inkonsistenz der Maße keinen Einfluss mehr hat (vgl. Abschnitt 2.4 Beispiel 2.2).

Weiters treffen die Maße Aussagen im Modalraum, wodurch aus den Ergebnissen nicht immer auf die Steuerbarkeit der physikalischen Größen geschlossen werden kann. Jedoch können sie wie am Beispiel „Doppelpendel“ zu sehen, einen guten Einblick in die einzelnen Systemteile liefern.

Das Maß κ_s nach LÜCKEL und MÜLLER erweist sich wegen der Kopplung zwischen Steuerbarkeit und Stabilität als nicht sinnvoll, da diese einerseits zu unerwünschten Verzerrungen führt und andererseits die beiden Systemeigenschaften prinzipiell von einander unabhängig sind. Die Polempfindlichkeitsmaße S, s nach LITZ und das Regelbarkeitsmaß ρ nach HIPPE liefern pro Eigenwert für jedes Eingangs-Ausgangspaar einen Wert. Dadurch erhält man mn Maßzahlen. Dies kann vor allem bei Mehrgrößensystemen höherer Ordnung schnell sehr unübersichtlich werden.

Die hier vorgestellten Energiemaße sind konsistent zu den KALMANSchen Begriffen. Der Nachteil dieser Maße ist die im Vergleich zu den Modalmaßen aufwendigere Berechnung, da die GRAMSche Matrix bestimmt werden muss. Das Invertieren der GRAMSchen Matrix \mathbf{Q}_S ist vor allem für Systeme höherer Ordnung aufwendig, und die Maßzahlen liefern keine zufriedenstellenden Ergebnisse, wenn die Matrix „nahezu singular“ ist. Der Vorteil dieser Maße besteht darin, dass direkt die Steuerbarkeit der physikalischen Zustandsvariablen bewertet wird. Da die Maße die Steuerbarkeit der Zustandsvariablen in den Ursprung bewerten, besteht auch hier eine Kopplung zwischen Steuerbarkeit und Stabilität.

Durch die Bestimmung von μ_C kann ein solcher Einblick ins Innere des Systems nicht erfolgen, dafür liefert dieses Maß einen Gesamteindruck über das Steuerbarkeitsverhalten des Systems. Die Anwendungsbeispiele zeigen, dass der Trisection-Algorithmus μ_C am zuverlässigsten bestimmt und durch die Toleranzschranke tol einstellbare und sehr enge Schranken liefern kann. Auch mittels der KRONECKER-Form können die obere und untere Schranke von μ_C zuverlässig bestimmt werden. Die Beispiele haben gezeigt, dass die KCF eine sehr gute obere Schranke liefert, während die untere Schranke in einigen Fällen deutlich kleiner ist als das Distanzmaß μ_C . Dieser Algorithmus hat jedoch gegenüber dem Trisection-Algorithmus den Vorteil, dass er um ein Vielfaches schneller ist (vgl. Tabelle 3.3). Die BFGS-Methode liefert ähnlich gute Schranken wie der Trisection-Algorithmus und ist um einiges schneller, kann aber im Falle ungünstiger Startwerte falsche Ergebnisse liefern. Die auf der Eigenwertabschätzung nach HE basierende Bestimmung von μ_C liefert im Allgemeinen auch gute Ergebnisse, hat aber im Falle von mehrfachen Eigenwerten Schwierigkeiten, da für diesen Fall die Modalform nicht definiert ist. Der Algorithmus von GAO und NEUMANN erweist sich als nicht sinnvoll, da er auf Grund der fix vorgegebenen Gitterpunkte ungenaue und nicht konsistente Ergebnisse liefert.

Wie bereits erwähnt ist für die Bestimmung von μ_R nur die reelle Achse des Pseudospektrums relevant. Dadurch ergeben sich für die Suche nach dem globalen Minimum einfachere Algorithmen, welche um ein Vielfaches schneller sind als die Algorithmen zur Bestimmung von μ_C . Am zuverlässigsten kann das Maß μ_R durch den von GAO und NEUMANN vorgestellten Algorithmus und durch die Nullstellensuche von $f(s)$ bestimmt werden.

Die in Abschnitt 3.2.2 vorgestellten oberen Schranken für μ_R spiegeln zwar in einigen Fällen den Verlauf der Änderung der Steuerbarkeit richtig wieder, liefern aber meist viel zu hohe Werte. Die Berechnung mittels Stufenform liefert in vielen Fällen konstante Ergebnisse und erweist sich deshalb als nicht sinnvoll, da in diesen Fällen keine Aussagen über den Steuerbarkeitsverlauf gemacht werden können. Auch die Bestimmung von μ_R mittels Quasi-KALMAN-Zerlegung liefert in einigen Fällen unbrauchbare Ergebnisse. Der Nachteile der Ermittlung des Distanzmaßes μ_R mit Hilfe der Formel nach GAHINET besteht darin, dass dafür der Wert μ_C benötigt wird. Auch wird für den Fall $\mu_C \neq \mu_R$ der Verlauf von μ_R nicht korrekt wiedergeben.

Für die einfachen Methoden zur Abschätzung des Distanzmaßes μ können die erhaltenen Ergebnisse wie folgt zusammengefasst werden: Bei allen in den vorherigen Abschnitten durchgeführten Beispielen liefert der kleinste Singulärwert σ_{\min} der Steuerbarkeitsmatrix \mathbf{S}_u Werte in der Größenordnung von μ_C , es gilt aber nicht $\mu_C \leq \sigma_{\min}$. I.A. gibt $\sigma_{\min}(\mathbf{S}_u)$ den Steuerbarkeitsverlust korrekt wieder. Wie im Beispiel aus der Einleitung gezeigt, kann es jedoch passieren das σ_{\min} auch für ein steuerbares System auf Grund von numerischen Ungenauigkeiten zu Null wird. Ein sehr gutes Ergebnis für die einfache Abschätzung des Distanzmaßes liefert die Schranke μ_λ . Diese Abschätzung stellt als einzige der vorgestellten einfachen Methoden eine obere Schranke dar und ist in vielen Fällen gleich dem Distanzmaß μ_C . Außerdem zeigt sie in den untersuchten Beispielen den Steuerbarkeitsverlust immer korrekt an.

Am Beispiel „Wagen mit Stab“ in Abschnitt 4.2.1.5 wurde gezeigt, dass durch die Linearisierung des Systems Informationen über die Steuerbarkeit des nichtlinearen Systems verloren gehen. Daher erscheint eine Maßdefinition die direkt die Steuerbarkeit eines nichtlinearen Systems quantifiziert sinnvoll und stellt ein interessantes Forschungsgebiet dar.

In den folgenden Abschnitten werden Bereiche vorgestellt, auf welche die Idee der Quantifizierung der Steuer- und Beobachtbarkeit angewendet werden kann. Wie die in dieser Arbeit untersuchten Maße für LZI-Systeme auf diese Bereiche adaptiert werden können, stellen weitere interessante Forschungsgebiete dar.

6.1 Eingangs- Ausgangssteuerbarkeit

Die in dieser Arbeit betrachtete Zustandssteuerbarkeit impliziert nicht, dass ein System auch von einem praktischen Gesichtspunkt aus steuerbar ist. Manche Systeme, die zwar theoretisch steuerbar sind, können in der Praxis auf Grund von physikalischen Grenzen wie z.B. Stellgrößenbeschränkungen, Messungenauigkeiten etc. dennoch nicht steuerbar sein. Umgekehrt gibt es auch Systeme, die im praktischen Sinne steuerbar sind, obwohl sie nicht zustandssteuerbar sind, da z.B. das Verhalten der nichtsteuerbaren Zustände außerhalb der Systemgrenzen liegt oder diese nicht von praktischer Bedeutung sind. Diese Steuerbarkeit im praktischen Sinne, die sogenannte Eingangs-Ausgangssteuerbarkeit, wird in [41] wie folgt definiert: Eingangs- Ausgangssteuerbarkeit ist die Fähigkeit, eine akzeptable Reglerperformance zu erreichen. D.h., die Ausganggröße y oder den Regelfehler e trotz unbekannter, aber begrenzter Ungenauigkeiten, wie z.B. Streckenveränderungen oder die Störung d , mit den zur Verfügung stehenden Eingangsgrößen u und Messgrößen y_m und d_m , in vorgegebenen Grenzen zu halten. In anderen Worten heißt das, dass die Strecke dann eingangs-ausgangssteuerbar ist, wenn ein Regler existiert, der es ermöglicht, die Strecke mit den vorhandenen Eingangsgrößen und Messgrößen und für alle zu erwartenden Streckenschwankungen und Störeinflüsse zu regeln.

Die in dieser Arbeit betrachteten Steuerbarkeitsmaße geben nur einen Einblick auf das Systemverhalten bezüglich der Zustandssteuerbarkeit, lassen aber die praktischen Gesichtspunkte außer Acht.

6.2 Perspektivische Beobachtbarkeit

In [17] wird der Begriff perspektivische Beobachtbarkeit erstmals allgemein eingeführt:

Definition 3 (Perspektivische Beobachtbarkeit): Ein lineares, zeitinvariantes System wird perspektivisch beobachtbar genannt, wenn aus der Kenntnis von $y(t)$ bis auf d Dimensionen in einem endlichen Zeitintervall $[0 T]$ der unbekannte Anfangszustand $x(0)$ bis auf d Dimensionen bestimmt werden kann.

Es sei p die Anzahl der Ausgänge des Systems. Dann sind im Gegensatz zur Beobachtbarkeit nur $p - d$ Dimensionen d.h. eine Projektion des Ausgangsvektors bekannt. Da nur diese Projektion zur Verfügung steht, ermittelt man eine Projektion des Anfangszustandes des Zustandsvektors auf $n - d$ Dimensionen.

Das Konzept der perspektivischen Beobachtbarkeit findet vor allem im Bereich des maschinellen Sehens Anwendung. Dort werden dreidimensionale Objekte der Umgebung auf einen zweidimensionalen Bildschirm projiziert, d.h., von einem 3-dimensionalen Vektor sind im Rechner nur noch $3 - 1$ Dimensionen bekannt. Da nur die Projektion des Ausgangsvektors bekannt ist, genügt es, eine Projektion des Anfangszustandes des Zustandsvektors auf $n - d$ Dimensionen zu ermitteln.

Einen weiteren Anwendungsfall für die perspektivische Beobachtbarkeit stellen Systeme mit einer rationalen Funktion am Ausgang dar. Der Ausgang

$$y = \frac{c_{11}x_1 + c_{12}x_2 + c_{13}x_3}{c_{21}x_1 + c_{22}x_2 + c_{23}x_3}$$

kann als Steigung des Vektors $[y_1 \ y_2]^T$ betrachtet werden:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \mathbf{C} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad \text{mit} \quad \mathbf{C} = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \end{bmatrix}.$$

D.h. man kennt die Projektion $y = \frac{y_1}{y_2}$. Perspektivische Beobachtbarkeit bedeutet in diesem Fall, dass aus der Kenntnis von $y(t)$ die Richtung bzw. die Projektion des Anfangszustandes ermittelt werden kann.

Zur Überprüfung der perspektivischen Beobachtbarkeit existieren dem HAUTUS-Kriterium ähnliche Kriterien. In [45] wird gezeigt, wie die Überprüfung von unendlich vielen Parametern auf eine endliche Menge reduziert und somit in der Praxis angewendet werden kann. Weiters wird erläutert, wie die Idee des Distanzmaßes auf die perspektivische Beobachtbarkeit erweitert werden kann, um auch diese quantitativ bewerten zu können.

Anhang A

Definitionen und Erläuterungen

A.1 Pseudospektrum

Es sei \mathbf{A} eine (n, n) -Matrix. Für die Eigenwerte von \mathbf{A} gilt

$$\Lambda(\mathbf{A}) = \{s \in \mathbb{C} : \det(s\mathbf{E} - \mathbf{A}) = 0\}.$$

Dies entspricht den Punkten s in der komplexen Ebene, für die $(s\mathbf{E} - \mathbf{A})^{-1}$ nicht definiert ist. An diesen Punkten strebt die Norm von $(s\mathbf{E} - \mathbf{A})^{-1}$ gegen Unendlich. Es ist nun auch von Interesse, das Verhalten von $\|(s\mathbf{E} - \mathbf{A})^{-1}\|$ in der Umgebung dieser Unstetigkeitsstellen zu betrachten. Das führt zu folgender Definition des Pseudospektrums [12]:

$$\Lambda_\epsilon(\mathbf{A}) := \{s \in \mathbb{C} : \|(s\mathbf{E} - \mathbf{A})^{-1}\| \geq \epsilon^{-1}\} \quad (\text{A.1})$$

für reelle $\epsilon > 0$. Das Pseudospektrum kann auch äquivalent zu (A.1) anhand der Eigenwerte von mit Störungen behafteten Matrizen definiert werden:

$$\Lambda_\epsilon(\mathbf{A}) := \{s \in \mathbb{C} : s \in \Lambda(\mathbf{A} + \delta\mathbf{A}) \quad \text{mit } \|\delta\mathbf{A}\| \geq \epsilon\}.$$

Zur Veranschaulichung wird als Beispiel das Pseudospektrum einer mit Zufallszahlen gefüllten $(50,50)$ -Matrix berechnet. Das Pseudospektrum wurde mit Hilfe von EigTool [52] in Matlab erstellt und ist in Abbildung A.1 dargestellt. Die reelle Achse ist in x- und die imaginäre Achse in y-Richtung aufgetragen. Die schwarzen Punkte stellen die Eigenwerte in der komplexen Ebene dar und die Farbskala zeigt die logarithmisch skalierten ϵ -Werte an. Abbildung A.1 kann folgendermaßen gedeutet werden: Wird die Matrix mit einer Störung der Norm 0,1 beaufschlagt, liegen die Eigenwerte der gestörten Matrix innerhalb der orangefarbenen Kurve. Deutet man die ϵ -Linien als Höhenschichtlinien, kann ein 3D-Gebirge erstellt werden, dessen Suprema den Eigenwerten entsprechen.

Für das Problem der Berechnung der Größe μ nach Relation 3.3 wird das Pseudospektrum ähnlich definiert:

$$\{s \in \mathbb{C} : \sigma_{\min}([s\mathbf{E} - \mathbf{A}, \mathbf{B}]) \leq \epsilon\}.$$

Hierbei handelt es sich nicht mehr um eine quadratische, sondern um eine rechteckige Matrix, deren Pseudospektrum weit weniger gut verstanden ist. Einige der Theoreme für das Pseudospektrum quadratischer Matrizen können zwar auf rechteckige Matrizen übertragen werden,

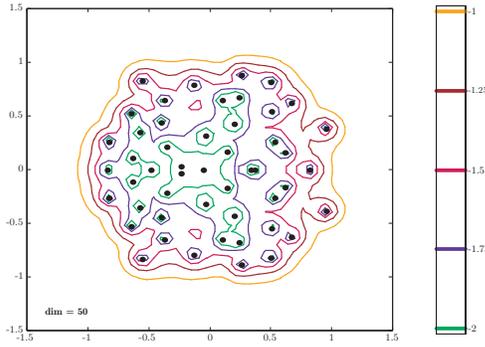


Abbildung A.1: Pseudospektrum einer (50, 50)-Matrix

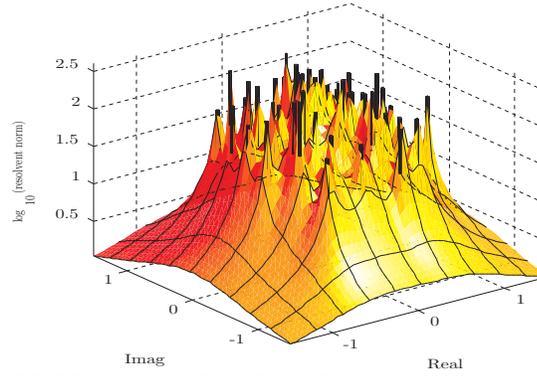


Abbildung A.2: Pseudospektrum einer (50, 50)-Matrix in 3D

jedoch unterscheiden sich die mathematischen Methoden bei der Untersuchung des Pseudospektrums rechteckiger Matrizen stark von denen der quadratischen. Beispielsweise ist für eine quadratische Matrix \mathbf{A} die Resolvente $(s\mathbf{E} - \mathbf{A})^{-1}$ eine analytische Funktion in s für $s \notin \Lambda(\mathbf{A})$, $\|(s\mathbf{E} - \mathbf{A})^{-1}\|$ erfüllt das Maximumprinzip und Projektionen und andere Matrizenoperationen können durch CAUCHY Integrale berechnet werden. All diese Eigenschaften treffen im rechteckigen Fall nicht mehr zu, was darauf zurückzuführen ist, dass die Pseudoinverse einer rechteckigen Matrix \mathbf{M} durch $\mathbf{M}^+ = (\mathbf{M}^H \mathbf{M})^{-1} \mathbf{M}^H$ berechnet wird [51].

A.2 KALMAN-Zerlegung

Die KALMAN-Zerlegung erfolgt durch eine reguläre Zustandstransformation $\mathbf{x} = \mathbf{Tz}$ eines Systems der Form

$$\frac{d\mathbf{x}}{dt} = \mathbf{Ax} + \mathbf{Bu}.$$

Der dadurch entstandene Vektor \mathbf{z} kann in vier Teilvektoren zerlegt werden

$$\mathbf{z} = [\mathbf{z}_{S\bar{B}}, \mathbf{z}_{SB}, \mathbf{z}_{\bar{S}\bar{B}}, \mathbf{z}_{\bar{S}B}]^T,$$

so dass

$\mathbf{z}_{S\bar{B}}$ die steuerbaren, aber nicht beobachtbaren Zustände,

\mathbf{z}_{SB} die steuer- und beobachtbaren Zustände,

$\mathbf{z}_{\bar{S}\bar{B}}$ die nicht steuer- und nicht beobachtbaren Zustände,

$\mathbf{z}_{\bar{S}B}$ die nicht steuerbaren, aber beobachtbaren Zustände,

enthält. Das transformierte System besitzt folgende Struktur:

$$\begin{aligned} \frac{d\mathbf{z}}{dt} &= \begin{bmatrix} \mathbf{A}_{S\bar{B}} & \mathbf{A}_{12} & \mathbf{A}_{13} & \mathbf{A}_{14} \\ \mathbf{0} & \mathbf{A}_{SB} & \mathbf{0} & \mathbf{A}_{24} \\ \mathbf{0} & \mathbf{0} & \mathbf{A}_{\bar{S}\bar{B}} & \mathbf{A}_{34} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{A}_{\bar{S}B} \end{bmatrix} \mathbf{z} + \begin{bmatrix} \mathbf{B}_{S\bar{B}} \\ \mathbf{B}_{SB} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} \mathbf{u} \\ \mathbf{y} &= [\mathbf{0} \quad \mathbf{C}_{SB} \quad \mathbf{0} \quad \mathbf{C}_{\bar{S}B}] \mathbf{z}. \end{aligned}$$

Die Teilsysteme $(\mathbf{A}_{S\bar{B}}, \mathbf{B}_{S\bar{B}})$ und $(\mathbf{A}_{SB}, \mathbf{B}_{SB})$ sind steuerbar, und die Teilsysteme $(\mathbf{A}_{\bar{S}B}, \mathbf{C}_{\bar{S}B})$ und $(\mathbf{A}_{SB}, \mathbf{C}_{SB})$ sind beobachtbar (vgl. Abbildung A.3).

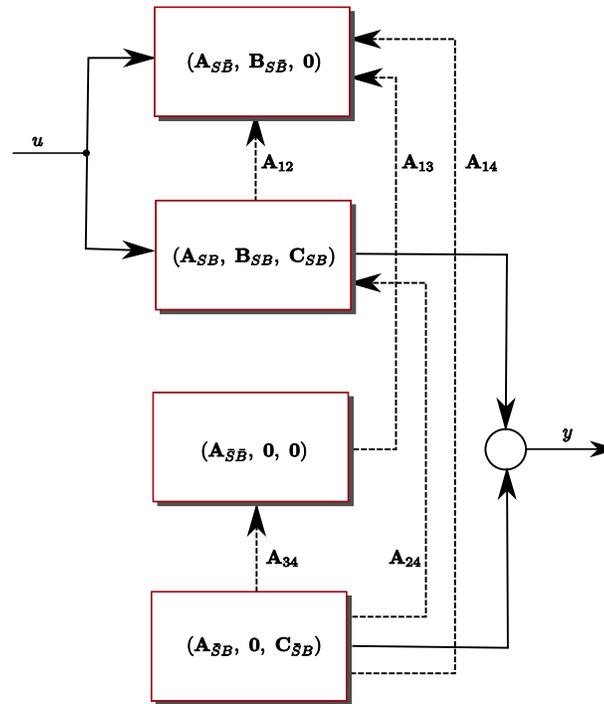


Abbildung A.3: KALMAN-Zerlegung

Durch die KALMAN-Zerlegung kann ein nicht steuerbares System in ein steuerbares und ein nicht steuerbares System aufgeteilt werden:

$$\begin{aligned} \frac{dz}{dt} &= \begin{bmatrix} \mathbf{A}_S & \mathbf{A}_{12} \\ \mathbf{0} & \mathbf{A}_{\bar{S}} \end{bmatrix} \mathbf{z} + \begin{bmatrix} \mathbf{B}_S \\ \mathbf{0} \end{bmatrix} \mathbf{u} \\ \mathbf{y} &= [\mathbf{C}_1 \quad \mathbf{C}_2]. \end{aligned}$$

Analoges gilt für ein nicht beobachtbares System.

A.3 KRONECKER kanonische Form

Durch die Transformation eines sogenannten *Matrizenbüschels* $\mathbf{F} + s\mathbf{G}$ in die KRONECKER kanonische Form

$$\mathbf{U}^{-1}(\mathbf{F} + s\mathbf{G})\mathbf{V} =: \mathbf{S} + s\mathbf{T}$$

erhält man eine Zerlegung in ein sogenanntes singuläres und reguläres Büschel.

Definition 4: Ein Matrizenbüschel $\mathbf{F} + s\mathbf{G}$ vom Typ (m', n') heißt *regulär*, wenn

1. \mathbf{F} und \mathbf{G} quadratische Matrizen n' -ter Ordnung sind und
2. die Determinante von $\mathbf{F} + s\mathbf{G}$ nicht identisch Null ist.

In jedem anderen Fall

1. $m' \neq n'$ oder
2. $m' = n'$, aber die Determinante des Büschels $\mathbf{F} + s\mathbf{G}$ ist identisch Null

heißt die Matrix $\mathbf{F} + s\mathbf{G}$ *singulär* [15].

Die beiden nun folgenden Sätze geben an, wie reguläre und singuläre Büschel weiter unterteilt werden können.

Lemma 13 (reguläre Büschel): Jedes reguläre Büschel $\mathbf{F} + s\mathbf{G}$ kann in die Form

$$\mathbf{U}^{-1}(\mathbf{F} + s\mathbf{G})\mathbf{V} = \mathbf{S} + s\mathbf{T} = \begin{bmatrix} \mathbf{N}_{u_1} & & & \\ & \mathbf{N}_{u_2} & & \mathbf{0} \\ & & \ddots & \\ \mathbf{0} & & & \mathbf{N}_{u_l} \\ & & & & \mathbf{J} + s\mathbf{E} \end{bmatrix}$$

transformiert werden. Die ersten l Diagonalelemente entsprechen den unendlichen Elementarteilern des Büschels $\mathbf{F} + s\mathbf{G}$. Das letzte Diagonalelement $\mathbf{J} + s\mathbf{E}$ ist durch die endlichen Elementarteiler¹ des betrachteten Büschels eindeutig definiert.

Die Matrix \mathbf{J} ist die sogenannte JORDAN kanonische Form, die folgende Struktur besitzt:

$$\mathbf{J} = \begin{bmatrix} \mathbf{J}_1 & & & \\ & \mathbf{J}_2 & & \mathbf{0} \\ & & \ddots & \\ & \mathbf{0} & & \mathbf{J}_k \end{bmatrix} \quad \text{mit} \quad \mathbf{J}_i = \begin{bmatrix} \lambda_i & 1 & & \\ & \lambda_i & 1 & \mathbf{0} \\ & & \ddots & \ddots \\ \mathbf{0} & & & \lambda_i & 1 \\ & & & & \lambda_i \end{bmatrix}.$$

Unendliche Elementarteiler und somit N Blöcke existieren genau dann, wenn die Determinante der Matrix \mathbf{G} gleich Null ist.

Es sei r der Rang des Büschels, d.h. die maximale Ordnung der nicht identisch verschwindenden Minoren. Für ein singuläres Büschel $\mathbf{F} + s\mathbf{G}$ vom Typ (m', n') gilt, dass stets mindestens eine der Ungleichungen $r < n'$ bzw. $r < m'$ erfüllt ist.

Gilt $r < n'$, dann sind die Spalten des Büschels linear abhängig, d.h. die Gleichung

$$(\mathbf{F} + s\mathbf{G})\mathbf{x} = \mathbf{0}$$

besitzt eine nichttriviale Lösung. Die folgenden Betrachtungen beschränken sich auf solche Lösungen $\mathbf{x}(s)$ von (A.3), die Polynome in s sind.

¹Für die Berechnung von Elementarteiler wird auf [15] verwiesen.

Lemma 14 (singuläre Büschel): Besitzt Gleichung (A.3) eine Lösung $\mathbf{x}(s)$ minimalen Grades ϵ_1 und ist $\epsilon_1 > 0$, so kann das vorgegebene Büschel $\mathbf{F} + s\mathbf{G}$ in ein Büschel der Form

$$\mathbf{S} + s\mathbf{T} = \begin{bmatrix} \mathbf{L}_{\epsilon_1} & \mathbf{0} \\ \mathbf{0} & \mathbf{F}_1 + s\mathbf{G}_1 \end{bmatrix}$$

mit der $(\epsilon_1, \epsilon_1 + 1)$ -Matrix

$$\mathbf{L}_{\epsilon_1} = \begin{bmatrix} s & 1 & & & \\ & s & 1 & & \\ & & \ddots & \ddots & \\ & & & s & 1 \\ \mathbf{0} & & & & \mathbf{1} \end{bmatrix}$$

überführt werden. Die Gleichung $(\mathbf{F}_1 + s\mathbf{G}_1)\mathbf{x} = \mathbf{0}$ besitzt keine Lösung, deren Grad kleiner als ϵ_1 ist.

Besitzt Gleichung $(\mathbf{F}_1 + s\mathbf{G}_1)\mathbf{x} = \mathbf{0}$ eine nichttriviale Lösung minimalen Grades ϵ_2 (dabei gilt $\epsilon_2 \geq \epsilon_1$), so kann durch Anwendung von Lemma 14 das Büschel $\mathbf{F}_1 + s\mathbf{G}_1$ erneut unterteilt werden. Durch wiederholtes Anwenden dieses Prozesses, wird das gegebene Büschel $\mathbf{F} + s\mathbf{G}$ in die Form

$$\mathbf{S} + s\mathbf{T} = \begin{bmatrix} \mathbf{L}_{\epsilon_1} & & & & \\ & \mathbf{L}_{\epsilon_2} & & & \mathbf{0} \\ & & \ddots & & \\ & \mathbf{0} & & \mathbf{L}_{\epsilon_p} & \\ & & & & \mathbf{F}_p + s\mathbf{G}_p \end{bmatrix} \quad r < n' \quad (\text{A.2})$$

überführt. Dabei gelten die Beziehungen $0 < \epsilon_1 \leq \epsilon_2 \leq \dots \leq \epsilon_p$. Die Gleichung $(\mathbf{F}_p + s\mathbf{G}_p)\mathbf{x} = \mathbf{0}$ besitzt keine nichttriviale Lösung.

Gilt $r < m'$, dann sind die Zeilen des Büschels $\mathbf{F} + s\mathbf{G}$ linear abhängig und das transponierte Büschel $\mathbf{F}^T + s\mathbf{G}^T$ kann in die Form (A.2) überführt werden. Die Werte $\epsilon_1, \epsilon_2, \dots, \epsilon_p$ werden in diesem Fall mit $0 < \eta_1 \leq \eta_2 \leq \dots \leq \eta_q$ bezeichnet. Man erhält

$$\mathbf{S} + s\mathbf{T} = \begin{bmatrix} \mathbf{L}_{\eta_1}^T & & & & \\ & \mathbf{L}_{\eta_2}^T & & & \mathbf{0} \\ & & \ddots & & \\ & \mathbf{0} & & \mathbf{L}_{\eta_p}^T & \\ & & & & \mathbf{F}_p + s\mathbf{G}_p \end{bmatrix} \quad r < m'.$$

Um die minimalen Grad des Polynoms $\mathbf{x}(s)$ aus Gleichung (A.3) und somit den Parameter ϵ_i bzw. η_i zu bestimmen, muss die Gleichung nicht gelöst werden, sondern kann mit Hilfe des folgenden Lemmas bestimmt werden:

Lemma 15: Die Zahl ϵ ist der kleinste Wert des Index k , für den in den Relationen

$$\rho_k \leq (k + 1)n$$

das Ungleichheitszeichen gilt.

Zusammenfassend kommt man zu dem Ergebnis, dass jedes (m', n') -Matrizenbüschel $\mathbf{F} + s\mathbf{G}$ in die KRONECKER kanonische Form $\mathbf{S} + s\mathbf{T}$, bestehend aus den Blöcken $\mathbf{L}^T, \mathbf{L}, \mathbf{N}$ und $\mathbf{J} + s\mathbf{E}$, transformiert werden kann (siehe Abbildung A.4).

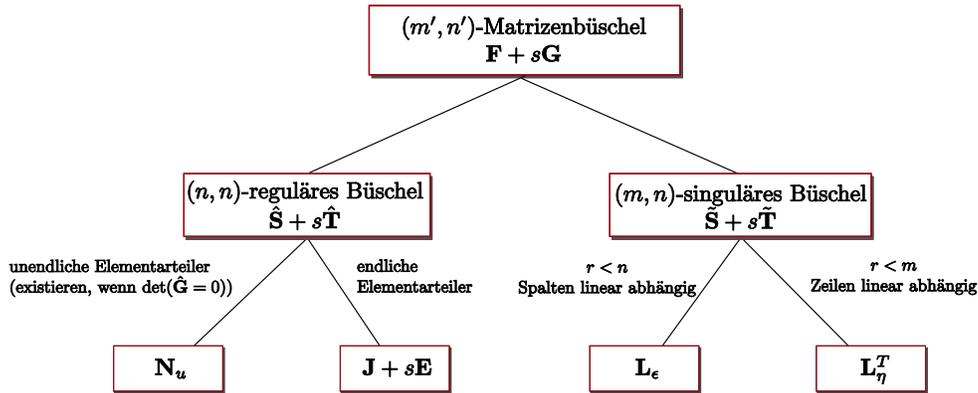


Abbildung A.4: Zerlegung in KRONECKER kanonische Form (KCF)

A.4 Störungstheorie

Die Störungstheorie beschäftigt sich mit der Frage, wie sich eine „Funktion“ verändert, wenn auf das Argument der Funktion eine Störung wirkt. Hierbei kann es sich bei der genannten Funktion um ein Polynom, eine Matrix oder auch ein Differentialgleichungssystem handeln. Das Ergebnis einer Störungsanalyse kann entweder eine Approximation der Störung oder eine obere Schranke für die Störung sein. Eine Störungsapproximation schätzt die Störung in Form von Funktionstermen in Abhängigkeit der bekannten Störung im Argument. Die Untersuchung von Schranken zur Störungsabschätzung startet mit einer bekannten Beschränkung der Störung im Argument und nutzt diese zur Ermittlung von Schranken für die Funktionswerte.

Betrachtet man die Definition des Distanzmaßes nach PAIGE (siehe Gleichung (3.2)), können die Matrizen $\delta\mathbf{A}$ und $\delta\mathbf{B}$ als Störung des Systems (\mathbf{A}, \mathbf{B}) aufgefasst werden. Wird das Distanzmaß μ von diesem Blickwinkel aus betrachtet, können die Erkenntnisse der Matrix Störungstheorie zur Ermittlung von Schranken für μ benutzt werden. In dieser Theorie wird eine Matrix \mathbf{A} betrachtet und untersucht, wie sich z.B. die Eigenwerte oder die Singulärwerte der gestörten Matrix $\tilde{\mathbf{A}} := \mathbf{A} + \Delta$ verhalten. Die für das Distanzmaß relevanten Theoreme der Störungstheorie werden in diesem Abschnitt erläutert (für ausführlichere Informationen siehe z.B. [3, 46, 50]).

A.4.1 Das BAUER-FIKE-Theorem

Das BAUER-FIKE-Theorem gibt eine Schranke für die Eigenwerte der gestörten Matrix $\tilde{\mathbf{A}} = \mathbf{A} + \Delta$ an.

Theorem 16 (BAUER-FIKE): Es sei \mathbf{Q} eine reguläre Matrix. Wenn $\tilde{\lambda}$ ein Eigenwert der Matrix $\tilde{\mathbf{A}}$ und kein Eigenwert der Matrix \mathbf{A} ist, dann gilt

$$\|\mathbf{Q}^{-1}(\mathbf{A} - \tilde{\lambda}\mathbf{E})^{-1}\mathbf{Q}\|^{-1} \leq \|\mathbf{Q}^{-1}\Delta\mathbf{Q}\|. \quad (\text{A.3})$$

Aus diesem Theorem lässt sich das vereinfachte Theorem für diagonalisierbare Matrizen direkt ableiten. Die Matrix \mathbf{A} sei diagonalisierbar, d.h. $\mathbf{X}_L\mathbf{A}\mathbf{X}_R = \Lambda$ mit $\mathbf{X}_L = \mathbf{X}_R^{-1}$ und $\Lambda = \text{diag}(\lambda_i)$. Durch Einsetzen von $\mathbf{Q} = \mathbf{X}_R$ in Gleichung (A.3) erhält man

$$\|(\Lambda - \tilde{\lambda}\mathbf{E})^{-1}\|^{-1} \leq \|\mathbf{X}_L\Delta\mathbf{X}_R\|.$$

Daraus ergibt sich für das BAUER-FIKE-Theorem:

Theorem 17 (BAUER-FIKE (für eine diagonalisierbare Matrix)): Es sei \mathbf{A} eine diagonalisierbare Matrix. Dann existiert für jeden Eigenwert $\tilde{\lambda} \in \lambda(\tilde{\mathbf{A}})$ ein Eigenwert $\lambda \in \lambda(\mathbf{A})$ so, dass gilt

$$|\lambda - \tilde{\lambda}| \leq K(\mathbf{X}_R) \|\Delta\|$$

mit der Konditionszahl $K(\mathbf{X}_R) = \|\mathbf{X}_R\| \|\mathbf{X}_L\|$.

Dieses Theorem liefert eine obere Schranke für den Abstand zwischen den Eigenwerten der Matrix \mathbf{A} und den Eigenwerten der gestörten Matrix $\tilde{\mathbf{A}} = \mathbf{A} + \Delta$. Diese ist nicht nur von der Störung Δ abhängig, sondern auch von der Konditionszahl $K(\mathbf{X}_R)$. Ist die Konditionszahl und somit die obere Schranke klein, spricht man von einem gut konditionierten Problem. Unitäre Matrizen haben die Konditionszahl Eins. Somit sind normale Matrizen gut konditioniert, da sich diese durch eine unitäre Transformation diagonalisieren lassen.

A.4.2 Das WIELANDT-HOFFMANN-Theorem und seine Konsequenzen

In diesem Abschnitt werden weitere Theoreme zur Abschätzung von Eigen- und Singulärwerten vorgestellt. Hierfür werden die Eigenwerte bzw. Singulärwerte folgendermaßen geordnet:

$$\operatorname{Re}\{\lambda_n\} \leq \dots \leq \operatorname{Re}\{\lambda_2\} \leq \operatorname{Re}\{\lambda_1\} \quad \text{bzw.} \quad \sigma_n \leq \dots \leq \sigma_2 \leq \sigma_1.$$

Theorem 18 (WIELANDT-HOFFMANN): Es seien \mathbf{A} und $\tilde{\mathbf{A}}$ normale Matrizen. Es existiert eine Permutation π ($\pi = 1, 2, \dots, n$) so, dass

$$\sum_{i=1}^n |\lambda_{\pi(i)} - \tilde{\lambda}_i|^2 \leq \|\tilde{\mathbf{A}} - \mathbf{A}\|_F^2 = \|\Delta\|_F^2 \quad \text{mit } i = 1, \dots, n$$

gilt.

D.h., die Eigenwerte der Matrix \mathbf{A} und die der gestörten Matrix $\tilde{\mathbf{A}}$ können so geordnet werden, dass die Quadratsumme der paarweisen Abstände kleiner ist als das Quadrat der Störung. Ist \mathbf{A} hermitesch, ergibt sich als Spezialfall des WIELANDT-HOFFMANN-Theorems folgende Abschätzung:

$$\sum_{i=1}^n |\lambda_i - \tilde{\lambda}_i|^2 \leq \|\tilde{\mathbf{A}} - \mathbf{A}\|_F^2 = \|\Delta\|_F^2 \quad \text{mit } i = 1, \dots, n.$$

Für die Abschätzung der Eigenwerte störbehafteter, unitärer Matrizen existiert nachfolgendes Theorem.

Theorem 19 (WEYL): Es seien \mathbf{A} und $\tilde{\mathbf{A}}$ unitäre Matrizen und $\|\cdot\|$ sei invariant gegenüber unitären Transformationen. Dann gilt die Ungleichung

$$\|\operatorname{diag}(|\lambda_i - \tilde{\lambda}_i|)\| \leq \|\Delta\|.$$

Für die Frobeniusnorm folgt obiges Theorem direkt aus dem WIELANDT-HOFFMANN-Theorem. WEYL hat gezeigt, dass es auch für die Spektralnorm gilt. MIRSKEY hat das Theorem für

alle gegenüber unitären Transformationen invarianten Normen bewiesen.

Aus dem WEYL-Theorem folgt die Ungleichung

$$|\lambda_i - \tilde{\lambda}_i| \leq \|\Delta\|.$$

Für die Abschätzung von Singulärwerten existiert folgendes, zum WEYL-Theorem analoges Theorem:

Theorem 20 (MIRSKY): Für alle Normen $\|\cdot\|$, die invariant gegenüber unitären Transformationen sind, gilt

$$\|\text{diag}(|\sigma_i - \tilde{\sigma}_i|)\| \leq \|\Delta\|.$$

Aus diesem Theorem folgt

$$|\sigma_i - \tilde{\sigma}_i| \leq \|\Delta\|.$$

Anhang B

Parameterwerte

Für die Beispiele „Hubmagnet“ und „Doppelpendel“ wurden folgende Parameterwerte angenommen:

Hubmagnet:

$$\begin{aligned}g &= 9,81 \text{ ms}^{-2} \\c &= 0,00013632 \text{ kgm}^3 \text{ s}^{-2} \text{ A}^{-2} \\m &= 0,06687 \text{ kg} \\L &= 1,08 \text{ H} \\R &= 18 \Omega\end{aligned}$$

Doppelpendel:

$$\begin{aligned}g &= 9,81 \text{ ms}^{-2} \\m_1 &= 1 \text{ kg} \\m_2 &= 2 \text{ kg} \\m_3 &= 10 \text{ kg} \\l_1 &= 2 \text{ m} \\l_2 &= 5 \text{ m} \\c &= 10 \text{ kgm}^{-1} \\k &= 1 \text{ Nm}^{-1}\end{aligned}$$

Literaturverzeichnis

- [1] BENNINGER, N. F.: *Analyse und Synthese linearer Systeme mit Hilfe neuer Strukturmaße*. VDI Verlag, 1987
- [2] BENNINGER, N. F. ; RIVOIR, J.: Ein neues konsistentes Maß zur Beurteilung der Steuerbarkeit in linearen, zeitinvarianten Systemen. In: *Automatisierungstechnik* 34 (1986)
- [3] BERNSTEIN, D.: *Matrix Mathematics*. Princeton University Press, 2005
- [4] BOLEY, D. L.: Estimating Distance to Uncontrollability - Upper and Lower Bounds. In: *Signal Processing, Scattering and Operator Theory, and Numerical Methods, MTNS-89* (1990), S. 373–380
- [5] BOLEY, D. L. ; LU, W.S.: Measuring How Far a Controllable System is from an Uncontrollable One. In: *IEEE Transactions on Automatic Control* AC-31 (1986), S. 249–251
- [6] BURKE, J. V. ; LEWIS, A. S. ; OVERTON, M. L.: Pseudospectral Components and the Distance to Uncontrollability. In: *SIAM Journal on Matrix Analysis and Applications* 26 (2005), Nr. 2, S. 350 – 361
- [7] DATTA, B. N.: *Numerical Methods for Linear Control Systems*. Elsevier Academic Press, 2004
- [8] DIETACHMAYR, F.: *Ermittlung der optimalen Abtastzeit am Labormodell 'Hubmagnet'*. – Bakkalaureatarbeit - TU Graz, Institut für Regelungs- und Automatisierungstechnik
- [9] EISING, R.: Between controllable and uncontrollable. In: *System & Control Letters* 4 (1984), S. 263–264
- [10] EISING, R.: The distance between a system and the set of uncontrollable systems. In: BEER-SHEVA, ed. P. A. F. P. A. Fuhrmann (Hrsg.): *Proceedings of the Mathematical Theory of Networks and Systems*. London : Springer-Verlag, 1984, S. 303–314
- [11] ELSNER, L. ; HE, C.: An algorithm for computing the distance to uncontrollability. In: *System & Control Letters* 17 (1991), Nr. 6, S. 453–464
- [12] EMBREE, M. ; TREFETHEN, L. N.: *Pseudospectra Gateway*. – <http://web.comlab.ox.ac.uk/projects/pseudospectra>
- [13] FÖLLINGER, O.: *Regelungstechnik*. Hüthig Buch Verlag Heidenberg, 1994
- [14] GAHINET, P.: Distance to the nearest uncontrollable pair and algebraic Riccati equation. In: *Proceedings of the 29th Conference on Decision and Control*, 1990, S. 273–278

- [15] GANTMACHER, F.: *Matrizentheorie*. Springer-Verlag, 1986
- [16] GAO, M. ; NEUMANN, M.: A global minimum search algorithm for estimating the distance to uncontrollability. In: *Linear Algebra and its Applications* 188/189 (1993), S. 305–350
- [17] GHOSH, B.K. ; LOUCKS, E.P.: An Introduction to Perspective Observability and Recursive Identification Problems in Machine Vision. In: *Proceedings of the 33rd IEEE Conference on Decision and Control* 4 (1994), S. 3229–3234
- [18] GU, M.: New Methods for Estimating the Distance to Uncontrollability. In: *SIAM Journal on Matrix Analysis and Applications* 21 (2000), Nr. 3, S. 989–1003
- [19] GU, M. ; MENGI, E. ; OVERTON, M.L. ; XIA, J. ; ZHU, J.: Fast Methods for Estimating the Distance to Uncontrollability. In: *SIAM Journal on Matrix Analysis and Applications* 28 (2006), Nr. 2, S. 477–502
- [20] HE, C.: Estimating the distance to uncontrollability: a fast method and a slow one. In: *System & Control Letters* 26 (1995), Nr. 4, S. 275–281
- [21] HIPPE, P.: Ein modales Regelbarkeitsmaß für lineare, zeitinvariante dynamische Systeme. In: *Regelungstechnik* 30 (1982), Nr. 3, S. 96 – 101
- [22] JOHNSON, C.: Optimization of a Certain Quality of Complete Controllability and Observability for Linear Dynamical Systems. In: *ASME Transactions Journal of Basic Engineering* 91 D (1969), S. 228 – 238
- [23] KALMAN, R. E. ; HO, Y. C. ; NARENDRA, K. S.: Controllability of Linear Dynamical Systems. In: *Contributions to Differential Equations* 1 (1963), Nr. 2, S. 189 – 213
- [24] KENNEY, C. ; LAUB, A. J.: Controllability and Stability Radii for Companion Form Systems. In: *Mathematics of Control, Signals, and Systems* 1 (1988), S. 239–256
- [25] KONNO, A. ; UCHIYAMA, M. ; KITO, Y. ; MURAKAMI, M.: Configuration-Dependent Controllability of Flexible Manipulators. In: *ISER*, 1993, S. 531–544
- [26] KRYSANDER, M. ; FRISK, E.: Sensor placement for fault diagnosis. In: *IEEE Transactions on Systems, Man, and Cybernetics – Part A: Systems and Humans* 38 (2008), S. 1398 – 1410
- [27] LÜCKEL, J. ; MÜLLER, P. C.: Analyse von Steuerbarkeits-, Beobachtbarkeits- und Störbarkeitsstrukturen linearer, zeitinvarianter Systeme. In: *Regelungstechnik* 23 (1975), S. 163 – 171
- [28] LEE, E. B. ; MARKUS, L.: *Foundations of Optimal Control Theory*. John Wiley & Sons, 1967 (The SIAM Series in Applied Mathematics)
- [29] LITZ, L.: Praktische Ergebnisse mit einem neuen modalen Verfahren zur Ordnungsreduktion. In: *Regelungstechnik* 27 (1979), S. 273 –280
- [30] LITZ, L.: Berechnung stabilisierender Ausgangsvektorrückführungen über Polempfindlichkeiten. In: *Regelungstechnik* 29 (1981), Nr. 12, S. 434 – 440

- [31] LITZ, L.: Dezentraler Regelerentwurf einer Durchlaufglühanlage mittels mehrstufiger Ordnungsreduktion. In: *Regelungstechnik* 30 (1982), S. 340 – 347
- [32] LITZ, L.: Modale Maße für Steuerbarkeit, Beobachtbarkeit, Regelbarkeit und Dominanz - Zusammenhänge, Schwachstellen, neue Wege. In: *Regelungstechnik* 31 (1983), Nr. 5, S. 148 – 158
- [33] LUNZE, J.: *Automatisierungstechnik*. Oldenbourg, 2003
- [34] MEHRMANN, V. ; XU, H.: An Analysis of the Pole Placement Problem. In: *Electronic Transactions on Numerical Analysis* 4 (1996), S. 89–105
- [35] MENGI, E.: *Measures for Robust Stability and Controllability*, New York University, Diss., 2006
- [36] MIMINIS, G. S. ; PAIGE, C. C.: An algorithm for pole assignment of time invariant linear systems. In: *International Journal of Control* 35 (1982), Nr. 2, S. 341–354
- [37] MÜLLER, P. C. ; WEBER, H. I.: Analysis and Optimization of Certain Qualities of Controllability and Observability for Linear Dynamical Systems. In: *Automatica* 8 (1972), S. 237–246
- [38] PAIGE, C. C.: Properties of numerical algorithms related to computing controllability. In: *IEEE Transactions on Automatic Control* 26 (1981), Nr. 1, S. 130 – 138
- [39] RIEGE, B.: Strukturmaße für dynamische Systeme / Universität Duisburg. 1998 (1). – Forschungsbericht
- [40] SEBASTIANI, I. ; TAMAS, M.: *Distanzmaß für zustandsgeregelte Systeme*. – Bakkalaureatarbeit - TU Graz, Institut für Regelungs- und Automatisierungstechnik
- [41] SKOGESTAD, S. ; POSTLETHWAITE, I.: *Multivariable Feedback Control: Analysis and Design*. John Wiley & Sons, 1996
- [42] SLOTINE, J. E. ; LI, W.: *Applied nonlinear control*. Prentice-Hall, 1991
- [43] SONTAG, E. D.: *Mathematical Control Theory*. Springer-Verlag, 1990
- [44] STADLER, M.: *Berechnung des Steuerbarkeitsmaßes nach Benninger*. 2009. – Projektarbeit - TU Graz, Institut für Regelungstechnik
- [45] STADLER, M.: *Überprüfung perspektivischer Beobachtbarkeit linearer zeitinvarianter Systeme*, TU Graz, Diplomarbeit, 2010
- [46] STEWART, G. W. ; SUN, J.: *Matrix Perturbation Theory*. Academic Press, Inc., 1990
- [47] STOER, J.: *Numerische Mathematik 1*. 8. Auflage. Springer Verlag, 1999
- [48] VAN DOOREN, P. M. ; GALLIVAN, K. A.: *State Space Analysis*. Vorlesungsunterlagen, 2006
- [49] WICKS, M. ; DECARLO, R.: Computing the Distance to an Uncontrollable System. In: *IEEE Transactions on Automatic Control* 36 (1991), S. 39–49
- [50] WILKINSON, J. H.: *The algebraic eigenvalue problem*. Oxford University Press, 1965

- [51] WRIGHT, T. ; TREFETHEN, L.: Pseudospectra of Rectangular Matrices. In: *IMA Journal of Numerical Analysis* 22 (2002), S. 501 – 519
- [52] WRIGHT, T.G.: *EigTool: A Graphical Tool for Nonsymmetric Eigenproblems*. <http://web.comlab.ox.ac.uk/pseudospectra/eigtool>, 2002
- [53] ZÖHRER, G.: *Fehlerdetektion und Isolation unter Zuhilfenahme des Steuerbarkeitsmaßes nach Benninger*. 2010. – Bakkalaureatarbeit - TU Graz, Institut für Regelungs- und Automatisierungstechnik